Multi- to one-dimensional transportation*

Pierre-André Chiappori[†], Robert J McCann[‡] and Brendan Pass[§]

Abstract

We consider the Monge-Kantorovich problem of transporting a probability density on $\mathbf{R}^m$ to another on the line, so as to optimize a given cost function. We introduce a nestedness criterion relating the cost to the densities, under which it becomes possible to solve this problem uniquely, by constructing an optimal map one level set at a time. This map is continuous if the target density has connected support. We use level-set dynamics to develop and quantify a local regularity theory for this map and the Kantorovich potentials solving the dual linear program. We identify obstructions to global regularity through examples.

More specifically, fix probability densities f and g on open sets $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}^n$ with $m \geq n \geq 1$. Consider transporting f onto g so as to minimize the cost $-s(x, y)$. We give a non-degeneracy condition (a) on $s \in C^{1,1}$ which ensures the set of

*The authors are grateful to Toronto's Fields' Institute for the Mathematical Sciences for its kind hospitality during part of this work. RJM acknowledges partial support of this research by Natural Sciences and Engineering Research Council of Canada Grant 217006-08 and -15, and by the US National Science Foundation under Grant No. DMS-14401140 while in residence at the Mathematical Sciences Research Institute in Berkeley CA during the spring 2016 semester. Chiappori gratefully acknowledges financial support from the NSF (Award 1124277). Pass is pleased to acknowledge support from Natural Sciences and Engineering Research Council of Canada Grant 412779-2012 and a University of Alberta start-up grant.

[†]Department of Economics, Columbia University, New York, USA pc2167@columbia.edu

[‡]Department of Mathematics, University of Toronto, Toronto, Ontario, Canada mccann@math.toronto.edu

[§]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada pass@ualberta.ca.

x paired with [g-a.e.] $y \in Y$ lie in a codimension n submanifold of X. Specializing to the case $m > n = 1$, we discover a nestedness criteria relating s to (f, g) which allows us to construct a unique optimal solution in the form of a map $F : X \longrightarrow \overline{Y}$. When $s \in C^2 \cap W^{3,1}$ and $\log f$ and $\log g$ are bounded, the Kantorovich dual potentials (u, v) satisfy $v \in C^{1,1}_{loc}(Y)$, and the normal velocity V of $F^{-1}(y)$ with respect to changes in y is given by $V(x) = v''(F(x)) - s_{yy}(x, F(x))$. Positivity (b) of V locally implies a Lipschitz bound on F; moreover, $v \in C^2$ if $F^{-1}(y)$ intersects $\partial X \in C^1$ transversally (c). On subsets where (a)-(c) can be be quantified, for each integer $r \geq 1$ the norms of $u, v \in C^{r+1,1}$ and $F \in C^{r,1}$ are controlled by these bounds, $\|\log f, \log g, \partial X\|_{C^{r-1,1}}$, $\|\partial X\|_{C^{1,1}}$, $\|s\|_{C^{r+1,1}}$, and the smallness of $F^{-1}(y)$. We give examples showing regularity extends from X to part of $\overline{X}$, but not from Y to $\overline{Y}$. We also show that when s remains nested for all (f, g), the problem in $\mathbf{R}^m \times \mathbf{R}$ reduces to a supermodular problem in $\mathbf{R} \times \mathbf{R}$.

1 Introduction

In the optimal transportation problem of Monge and Kantorovich, one is provided with probability measures $d\mu(x)$ and $d\nu(y)$, and asked to couple them together so as to minimize a given transportation cost, or equivalently to maximize a given surplus function $s(x, y)$. The measures are defined on subsets X and Y of complete separable metric spaces, often Euclidean spaces or manifolds with additional structure, with the surplus function $s(x, y)$ either defined by or defining the geometry of the product $X \times Y$. Such problems have a wealth of applications ranging from the pure mathematics of inequalities, geometry and partial differential equations to topics in computer vision, design, meteorology, and economics. These are surveyed in the books of Rachev and Rüschendorf [42], Villani [50] [51], Santambrogio [46] and Galichon [18]. It has most frequently been studied under the assumption that $X = Y$, as in Monge [40] and Kantorovich [26], or at least that the two spaces X and Y have the same finite dimension. Monge's question concerned solutions in the form of maps $F : X \longrightarrow Y$ carrying μ onto ν. At the writing of [50], Villani described the regularity of such maps as the major open problem in the subject. Through the work of many authors, an intricate theory has been developed for the case of equal dimensions, leading up to the restrictive conditions of Ma, Trudinger and Wang [34] [49] under which the solution concentrates on the graph of a *smooth* map between X and Y; see [13] and [37] for complementary surveys.

Despite its relevance to applications, much less is known when $m \neq n$. The purpose of the present paper is to resolve this situation, at least in the case $m > n = 1$ motivating our title. Though it seems not to have received much attention previously, we view this as a case interpolating between between $m = n = 1$, which can be solved exactly in the supermodular case $\partial^2 s/\partial x \partial y > 0$, and the fully general problem $m \geq n > 1$. We develop, for the first time, a regularity theory addressing this intermediate case. It is based crucially on a *nestedness* condition introduced simultaneously here and in our companion work [8]. When satisfied, we show this condition leads to a unique solution in the form of a Monge mapping. This solution, moreover is semi-explicit: it is based on identifying each level set of the optimal map independently, and can presumably be computed with an algorithmic complexity significantly smaller than non-nested problems of the same dimensions can be solved. Although nestedness is restrictive, we suspect it may be a requirement for the continuity of optimal maps. Moreover, it depends subtly on the relation between μ, ν and s, which distinguishes it sharply from the familiar criteria for mappings, uniqueness and regularity when $m = n$ which, with few exceptions [27], depend primarily on the geometry and topology of s. Our theory addresses interior regularity, as well as regularity at some parts of the boundary. We identify various obstacles to nestedness and regularity along the way, including examples which show that higher regularity cannot generally hold on the entire boundary.

The plan for the paper is as follows. The next section introduces notation, describes the problem more precisely, and recalls some of its history and related developments. It is followed by a section dealing with general source and target dimensions $m \geq n \geq 1$, giving conditions under which the set of x paired with a.e. $y \in Y$ lie in a codimension n submanifold of X. In section §4 we specialize to $n = 1$, introduce the notion of nestedness, and show that it allows us to obtain a unique solution in the form of a optimal map between μ and ν which, under suitable conditions, is continuous. The solution map $F : X \longrightarrow Y$ is constructed one level set at a time. Although nestedness generally depends on the relation of (μ, ν) to s, in section §4.2 we show that when it happens to hold for all absolutely continuous (μ, ν), then s can effectively be reduced to a super- or submodular function of two real variables y and $I(x)$. In Section §5 we explore the motion of the level sets $F^{-1}(y)$ as $y \in Y$ is varied. To do so requires additional regularity of the data (s, μ, ν), under which we are able to deduce some additional regularity of the solution and give several conditions equivalent to nestedness. In Section §6 we give examples of nested and non-nested problems, including some which illustrate why the unequal dimensions of the problem prevent F from extending smoothly to the entire boundary of X. Finally, in Section 8, we develop a complete theory which describes how the higher regularity of F (and the Kantorovich dual potentials) is controlled by certain parameters governing (s, μ, ν, X, Y) and various geometrical aspects of the problem, such as the smallness of $F^{-1}(y)$, its transversality to ∂X, and the speed of its motion, which quantifies the uniformity of nestedness by determining the separation between distinct level sets.

Our interest in this problem was initially motivated by economic matching problems with transferable utility [8], such as the stable marriage problem [29], in which μ represents the distribution of female and ν the distribution of male types, and $s(x, y)$ represents the marital surplus obtained, to be divided competitively between a husband of type y and a wife of type x. The equivalence of optimal transportation to stable matching with transferable utility was shown in the discrete setting by Shapley and Shubik [47]. We adopt this terminology hereafter, in spite of the fact that μ and ν might equally well represent producer and consumer locations, buyer and seller preferences, etc. There is no reason a priori to expect the characteristics describing wives and husbands to have the same dimension. Although we anticipate that this theory will have many other applications, we are particularly aware of its potential for use in the semigeostrophic model of atmosphere and ocean dynamics, in which μ would represent the distribution of fluid in the physical domain and ν a potential vorticity sheet or filament in dual coordinates [11].

2 Setting and background results

Given Borel probability measures μ on $X \subset \mathbf{R}^m$ and ν on $Y \subset \mathbf{R}^n$, the Monge-Kantorovich problem is to transport μ onto ν so as to optimize the given surplus function $s(x, y)$. From Theorem 4 onwards, we assume X and Y are open, but for the moment they remain arbitrary. Assuming $s \in C(X \times Y)$ to be bounded and continuous for simplicity, we seek a Borel measure $\gamma \geq 0$ on $X \times Y$ having μ and ν for its marginals. The set $\Gamma(\mu, \nu)$ of such γ is convex, and weak-$*$ compact in the Banach space dual to $(C(X \times Y), \|\cdot\|_\infty)$. Among such γ, Kantorovich's problem is to maximize the linear functional

$$MK^* := \max_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} s(x, y) d\gamma(x, y). \tag{1}$$

We shall also be interested in the structure of the optimizer(s) γ. For example, is there a map $F : X \longrightarrow Y$ such that γ vanishes outside $Graph(F)$, and if so, what can be said about its analytical and geometric properties? Such a map is called a *Monge* or *pure* solution, *deterministic coupling*, *matching function* or *optimal map*, and we have $\gamma = (id \times F)_{\#}\mu$ where id denotes the identity map on X in that case [1]. More generally, if $F : X \longrightarrow Y$ is any μ-measurable map, we define the push-forward $F_{\#}\mu$ of μ through F by

$$F_{\#}\mu(V) = \mu[F^{-1}(V)]$$

for each Borel $V \subset Y$. Thus $\Gamma(\mu, \nu) = \{\gamma \geq 0 \text{ on } X \times Y \mid \pi^X_{\#}\gamma = \mu \text{ and } \pi^Y_{\#}\gamma = \nu\}$, where $\pi^X(x, y) = x$ and $\pi^Y(x, y) = y$. We denote by $\overline{X}$ the closure of X, and by $\operatorname{spt}\mu \subset \overline{X}$ the smallest closed set carrying the full mass of μ.

The case which is best understood is the case $X = Y = \mathbf{R}$, so that $n = m = 1$. If $s \in C^2(\mathbf{R}^2)$ satisfies

$$\frac{\partial^2 s}{\partial x \partial y}(x, y) > 0 \tag{2}$$

for all $x, y \in \mathbf{R}$, then (1) has a unique solution; moreover, this solution coincides with the unique measure $\gamma \in \Gamma(\mu, \nu)$ having non-decreasing support, meaning $(x, y), (x', y') \in \operatorname{spt}\gamma$ implies $(x - x')(y - y') \geq 0$. This result, which dates back to Lorentz [33], was rediscovered by the economists Becker [3], Mirrlees [39] and Spence [48]. Condition (2) is also called the Spence-Mirrlees condition, or *supermodularity*, in the economics literature. Note this condition does not depend on μ or ν, and while γ depends on them, it is independent of s in this case. Moreover, if μ is free from atoms, then γ concentrates on the graph of a non-decreasing function $F : \mathbf{R} \longrightarrow \mathbf{R}$ given by

$$\int_{(-\infty, F(x))} d\nu \leq \int_{-\infty}^{x} d\mu \leq \int_{(-\infty, F(x)]} d\nu. \tag{3}$$

When μ and ν are given by L^1 probability densities f and g, the fundamental theorem of calculus yields a ordinary differential equation

$$f(x) = F'(x) g(F(x)) \tag{4}$$

satisfied Lebesgue almost everywhere; smoothness properties of $F(x)$ can then be deduced from those of f and g.

In higher equal dimensions $m = n > 1$ the situation is much more subtle. However if μ is given by a probability density $f \in L^1(\mathbf{R}^m)$, it is again possible to given conditions on the surplus function $s \in C^1(X \times Y)$ such that the Kantorovich optimizer (1) is unique [38] [7] and concentrated on the graph of a map F [19] [30] depending sensitively on the choice of surplus. If ν is also given by a probability density $g \in L^1(\mathbf{R}^n)$, and $\log f, \log g \in L^\infty$, then it is possible to give conditions on $s \in C^4(X \times Y)$ which guarantee F is Hölder continuous [32] [16], and inherits higher regularity from that of f and g [31].

For unequal dimensions $m \geq n$, the existence, uniqueness and graphical structure of solutions follows from conditions on s as when $m = n$, but concerning other aspects of the problem much less is known. Only a result of Pass asserts that if smoothness of F holds for all $\log f, \log g \in C^\infty$, then the dimensions are effectively equal, in the sense that there are functions $I : \mathbf{R}^m \to \mathbf{R}^n$, $\alpha \in C(\mathbf{R}^m)$ and $\sigma \in C(\mathbf{R}^{2n})$ such that $s(x, y) = \sigma(I(x), y) + \alpha(x)$. If $n = 1$ we call I an *index* and s *pseudo-index* in this case; otherwise I represents a set of indices and s is *pseudo-indicial*.

In general, one of the keys to understanding the Kantorovich problem (1) is the dual linear program

$$MK_* := \inf_{(u,v)\in Lip_s} \int_X u(x)d\mu(x) + \int_Y v(y)d\nu(y), \tag{5}$$

where Lip_s consists of all pairs of payoff functions $(u, v) \in L^1(\mu) \oplus L^1(\nu)$ satisfying the stability constraint

$$u(x) + v(y) - s(x, y) \geq 0 \tag{6}$$

on $X \times Y$. The remarkable fact is that $MK_* = MK^*$ [26]. Thus, if γ and (u, v) optimize their respective problems, it follows that γ vanishes outside the zero set S of the non-negative function $u+v-s$. We therefore obtain the first and second order conditions

$$(Du(x), Dv(y)) = (D_x s(x, y), D_y s(x, y)) \tag{7}$$

and

$$\begin{pmatrix} D^2u(x) & 0 \\ 0 & D^2v(y) \end{pmatrix} \geq \begin{pmatrix} D^2_{xx}s(x,y) & D^2_{xy}s(x,y) \\ D^2_{yx}s(x,y) & D^2_{yy}s(x,y) \end{pmatrix} \tag{8}$$

at each $(x, y) \in S \cap (X \times Y)^0$ for which the derivatives in question exist; here X^0 denotes the interior of X. When the surplus s is Lipschitz, then the infimum (5) is attained by a pair of Lipschitz functions (u, v); when s has Lipschitz derivatives, we may even take u, v to be semiconvex, hence twice differentiable (in the sense of having a second-order Taylor expansion) Lebesgue almost everywhere [50] [46]. We let $\mathrm{Dom}\, Dv$ and $\mathrm{Dom}\, D^2v$ denote the domains where v admits a first- and second-order Taylor expansion, with $\mathrm{Dom}_0 Dv = \left(\overline{Y}\right)^0 \cap \mathrm{Dom}\, Dv$ and $\mathrm{Dom}_0 D^2v := \left(\overline{Y}\right)^0 \cap \mathrm{Dom}\, D^2v$ and $\overline{Y}$ denoting the closure of Y.

For each integer $r \geq 0$, and Hölder exponent $0 < \alpha \leq 1$ we denote by $C^{r,\alpha}(X)$ the space of functions which are r times continuously differentiable, and whose r-th derivatives are all Lipschitz continuous functions with respect to the distance function $|x - x'|^\alpha$ on X (in which case both properties extend to the closure $\overline{X}$ of X.) We norm this space by

$$\|f\|_{C^{r,\alpha}(X)} := \sum_{i=0}^{r} \sum_{|\beta|=i} \|D^\beta f\|_\infty + \sup_{x \neq x' \in X} \sum_{|\beta|=r} \frac{|D^\beta f(x') - D^\beta f(x)|}{|x' - x|^\alpha}$$

where $D^\beta f = \frac{\partial^{|i|} f}{\partial x_1 \cdots \partial x_i}$ and the sums are over multi-indices β of degree $|\beta|$. We eliminate the supremum if $\alpha = 0$, abbreviating $C^r(X) := C^{r,0}(X)$, $C(X) := C^0(X)$, and defining local versions $C^{r,\alpha}_{loc}(X)$ of these spaces analogously. We adopt the convention that $C^{-1,1} = L^\infty$, and denote by $\|\partial X\|_{C^{r,1}(X')}$ a minimal bound for the sums of norms of the $C^{r,1}$ functions parameterizing $\overline{X'} \cap \partial X$. In case $\alpha = 1$, this space forms an algebra; moreover, we extend these definitions to tensor fields f on Riemannian manifolds X by requiring the components of f to belong to $C^{r,1}_{loc}$ and setting

$$\|f\|_{C^{r,1}(X)} := \sum_{i=0}^{r+1} \|D^i f\|_\infty,$$

the operator D^i now denoting iterated covariant derivatives with respect to the Levi-Civita connection.

We shall also use the Sobolev spaces $W^{r,1}(X)$, consisting of integrable functions on X whose distributional partial derivatives up to order r are also integrable. This space is normed by

$$\|f\|_{W^{r,1}(X)} := \int_X (\sum_{|\beta| \le r} |D^\beta f|) d\mathcal{H}^m,$$

where the sum is over multi-indices β and $\mathcal{H}^m$ denotes the Hausdorff m-dimensional measure on X. Equipped with the norm

$$\|f\|_{(C \cap W^{1,1})(X)} := \max\{\|f\|_{L^\infty(X)}, \|f\|_{W^{1,1}(X)}\},$$

the space $(C \cap W^{1,1})(X)$ forms an algebra (closed under the continuous operation of multiplication), as does $C(Y; (C \cap W^{1,1})(X))$.

3 Transportation between unequal dimensions

We now turn our attention to the case in which the source and target domains X and Y of our transportation problem have unequal dimensions $m \ge n$. In this case, one expects many-to-one rather than one-to-one matching. Indeed, it is natural to expect that at equilibrium the subset $F^{-1}(y) \subset X \subset \mathbf{R}^m$ of partners which a man of type $y \in \mathrm{Dom}_0\, Dv$ is indifferent to will generically have dimension $m - n$, or equivalently, codimension n. Our first proposition recalls from [8] conditions under which this indifference set will in fact be a Lipschitz (or smoother) submanifold of the expected dimension. It is stated here for costs $s \notin C^2$ which need not be quite as smooth as those of [8].

3.1 Potential indifference sets

For any optimal γ and payoffs (u, v), we have already seen that $(x, y) \in S \cap (X \times \mathrm{Dom}_0 Dv)$ implies

$$D_y s(x, y) = Dv(y). \tag{9}$$

That is, all partner types $x \in X$ for husband $y \in \mathrm{Dom}_0 Dv$ lie in the same level set of the map $x \mapsto D_y s(x, y)$. If we know $Dv(y)$, we can determine this level set precisely; it depends on μ and ν as well as s. However, in the absence of this knowledge it is useful to define the *potential indifference sets*, which for given $y \in Y$ are merely the level sets of the map $x \in X \mapsto D_y s(x, y)$. We can parameterize these level sets by (cotangent) vectors $k \in T_y^* Y = \mathbf{R}^n$:

$$X(y, k) := \{x \in X \mid D_y s(x, y) = k\}, \tag{10}$$

or we can think of $y \in Y$ as inducing an equivalence relation between points of X, under which x and $\bar{x} \in X$ are equivalent if and only if

$$D_y s(x, y) = D_y s(\bar{x}, y).$$

Under this equivalence relation, the equivalent classes take the form (10). We call these equivalence classes *potential indifference sets*, since they represent a set of partner types which $y \in \mathrm{Dom}_0 Dv$ has the potential to be indifferent between.

A key observation concerning potential indifference sets is the following proposition. Recall for a Lipschitz function $F : \mathbf{R}^m \longrightarrow \mathbf{R}^n$, the *generalized Jacobian* or *Clarke subdifferential* $\partial F(x)$ at $x \in \mathbf{R}^m$ consists of the convex hull of limits of the derivatives of F at nearby points of differentiability [9]. For example, $\partial F(x) = \{DF(x)\}$ if F is C^1 at x.

Definition 1 (Surplus degeneracy) *Given $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}^n$ with $m \geq n$, we say $s \in C^2(X \times Y)$* degenerates *at $(\bar{x}, \bar{y}) \in X \times Y$ if $\mathrm{rank}(D^2_{xy} s(\bar{x}, \bar{y})) < n$. If $s \in C^{0,1}_{loc}(X \times Y)$ and $D_y s$ is locally Lipschitz, we say s degenerates at $(\bar{x}, \bar{y}) \in X \times Y$ if for every Lipschitz extension of $D_y s$ to a neighbourhood of $(\bar{x}, \bar{y})$ and choice of orthonormal basis for $\mathbf{R}^m$, setting*

$$F(x) = D_y s(x, \bar{y}) \tag{11}$$

yields some $m \times n$ matrix $M \in \partial F(\bar{x})$ with $\det[M_{ij}]_{1 \leq i,j \leq n} = 0$. Otherwise we say s is non-degenerate *at $(\bar{x}, \bar{y})$.*

Proposition 2 (Structure of potential indifference sets) *Let $s, D_y s \in C^{r,1}_{loc}(X \times Y)$ for some $r \geq 0$, where $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}^n$ with $m \geq n$. If s does not degenerate at $(\bar{x}, \bar{y}) \in X \times Y$, then $\bar{x}$ admits a neighbourhood $U \subset \mathbf{R}^m$ such that $X(\bar{y}, D_y s(\bar{x}, \bar{y})) \cap U$ coincides with the intersection of X with a $C^{r,1}$-smooth, codimension n submanifold of U.*

Proof. For smooth s, the set $L := \{x \in U \mid D_y s(x, \bar{y}) = D_y s(\bar{x}, \bar{y})\}$ forms a codimension n submanifold of U, by the preimage theorem [23, §1.4]. Otherwise choose an $C^{r,1}$ extension of $D_y s$ to a neighbourhood $U \times V$ of $(\bar{x}, \bar{y})$ and an orthonormal basis for $\mathbf{R}^m$ such that each $M \in \partial F$ has full rank with F from (11). The Clarke inverse function theorem [9] gives a $C^{r,1}$ local inverse to the map $x \in U \mapsto (D_y s(x, \bar{y}), x_{n+1}, \ldots, x_m)$. Taking U smaller if necessary, the set $X(\bar{y}, D_y s(\bar{x}, \bar{y}))$ is the image of the affine subspace $\{D_y s(\bar{x}, \bar{y})\} \times \mathbf{R}^{m-n}$ under this (biLipschitz) local inverse. ■

Although we have stated the proposition in local form, it implies that if $\bar{k} = D_y s(\bar{x}, \bar{y})$ is a *regular value* of $x \in X \mapsto D_y s(x, \bar{y})$ — meaning s is non-degenerate throughout $X(\bar{y}, \bar{k})$ — then $X(\bar{y}, \bar{k})$ is the intersection of X with an $m - n$ dimensional submanifold of $\mathbf{R}^m$.

Remark 3 (Genericity) *For $s \in C^2$ regular values are generic in the sense that for any given $\bar{y} \in \overline{Y}$, Sard's theorem asserts that the regular values of $D_y s(\cdot, \bar{y})$ form a set having full Lebesgue measure in $\mathbf{R}^n$. However, if $D_y s(\cdot, \bar{y})$ also has* critical *(i.e. non-regular) values for some $\bar{y} \in \overline{Y}$, it is entirely possible that $Dv(\bar{y})$ is a critical value of $D_y s(\cdot, \bar{y})$ for each such $\bar{y}$. This is necessarily the case when rank$(D^2_{xy} s(x, y)) <$ $\min\{m, n\}$ throughout $X \times Y$, meaning s is globally degenerate.*

As argued above, the potential indifference sets (10) are determined by the surplus function $s(x, y)$ without reference to the populations μ and ν to be matched. On the other hand, the indifference set actually realized by each $y \in Y$ depends on the relationship between μ, ν and s. This dependency is generally complicated. However, there is one case in which it may simplify substantially: the case of multi-to-one dimensional matching, namely $n = 1$. In this case, suppose $D^2_{xy} s(\cdot, y)$ is non-vanishing (i.e. $\frac{\partial s}{\partial y}(\cdot, y)$ takes only regular values). Then the potential indifference sets $X(y, k)$ form hypersurfaces in $\mathbf{R}^m$. Moreover, as k moves through $\mathbf{R}$, these potential indifference sets sweep out more and more of the mass of μ. For each $y \in Y$ there will be some choice of $k \in \mathbf{R}$ for which the μ measure of $\{x \mid D_y s(x, y) \leq k\}$ exactly coincides with the ν measure of $(-\infty, y]$ (assuming both measures are absolutely continuous with respect to Lebesgue, or at least that μ concentrates no mass on hypersurfaces and ν has no atoms). In this case the potential indifference set $X(y,k)$ is said to split the population proportionately at y, making it a natural candidate for being the *iso-husband* set $F^{-1}(y)$ to be matched with y.[1] In the next sections, we go on to describe and contrast situations in which this expectation is born out and leads to a complete solution from those in which it does not.

4 Multi-to-one dimensional matching

We now detail a new approach to a specific class of transportation problems, largely unexplored, but which can often be solved explicitly as below. These are *multi-to-one dimensional* problems, in which the space of wives may have several dimensions but the space of husbands only one. Thus, we are matching a distribution on $x = (x_1, ..., x_m) \in \mathbf{R}^m$ with another on $y \in \mathbf{R}$. The surplus s is then a function $s(x_1, ..., x_m, y)$ of $m+1$ real variables.

The goal is to construct from data (s, μ, ν) a mapping $F : X \longrightarrow Y \subset \mathbf{R}$, whose level sets $F^{-1}(y)$ constitute the *iso-husband* sets, or submanifold of wives among which husband x turns out to be indifferent facing the given market conditions. At the end of the preceding section we identified a natural candidate for this iso-husband set: namely the potential indifference set which divides the mass of μ in the same ratio as y divides ν; whether or not these natural candidates actually fit together to form the level sets of a function F or not depends on a subtle interaction between μ, ν and s. When they do, we say the model is *nested*, and in that case we show that the resulting function $F : X \longrightarrow Y$ produces the unique optimizer $\gamma = (id \times F)_{\#}\mu$ for (1). Note that except in the Lorentz/Becker/Mirrlees/Spence case $m = 1 = n$, this nestedness depends not only on s, but also on μ and ν.

4.1 Constructing explicit solutions for nested data

For each fixed $y \in Y \subseteq \mathbf{R}$, our goal is to identify the iso-husband set $\{x \in X \mid F(x) = y\}$ of husband type y in the given problem. When differentiability of v holds at $y \in Y^0$, the argument in the preceding section implies that this is contained in one of the potential indifference sets $X(y,k)$ from (10). Proposition 2 indicates when this set will have codimension 1; it generally divides X into two pieces: the sublevel set

$$X_{\leq}(y,k) := \{x \in X \mid \frac{\partial s}{\partial y}(x,y) \leq k\}, \tag{12}$$

and its complement $X_>(y,k) := X \setminus X_{\leq}(y,k)$. We denote its strict variant by $X_<(y,k) := X_{\leq}(y,k) \setminus X(y,k)$.

[1]Since $k = s_y(x,y)$ can be recovered from any $x \in X(y,k)$ and y, we may equivalently say x splits the population proportionately at y, and vice versa.

To select the appropriate level set, we choose the unique level set *splitting the population proportionately* with y; that is, the $k = k(y)$ for which the μ measure of female types $X_{\leq}(y,k)$ coincides with the ν measure of male types $(-\infty, y]$. We then set $y := F(x)$ for each x in $X(y,k)$. Our next theorem specifies conditions under which the resulting match $\gamma = (id \times F)_{\#}\mu$ optimizes the Kantorovich problem (1); we view it as the natural generalization of the positive assortative matching results of [33] [39] [3] and [48] from the one-dimensional to the multi-to-one dimensional setting. Unlike their criterion (2), which depends only on s, ours relates s to μ and ν, by requiring the sublevel sets $y \in Y \mapsto X_{\leq}(y, k(y))$ identified by the procedure above to depend monotonically on $y \in \mathbf{R}$, with the strict inclusion $X_{\leq}(y,k(y)) \subset X_{<}(y', k(y'))$ holding whenever $\nu[(y,y')] > 0$. We say the model (s,μ,ν) is *nested* in this case.

Theorem 4 (Optimality of nested matchings) *Let $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}$ be connected open sets equipped with Borel probability measures μ and ν. Assume ν has no atoms and μ vanishes on each Lipschitz hypersurface. Use $s \in C^{1,1}(X \times Y)$ and $s_y = \frac{\partial s}{\partial y}$ to define $X_{\leq}$, $X_{<}$, etc. as in* (12)*.*

(a) Assume s is non-degenerate throughout $X \times Y$. Then for each $y \in Y$ there is a maximal interval $K(y) = [k^-(y), k^+(y)] \neq \emptyset$ such that $\mu[X_{\leq}(y,k)] = \nu[(-\infty, y)]$ for all $k \in K(y)$. Both k^+ and $(-k^-)$ are upper semicontinuous.

(b) In addition, assume both maps $y \in Y \mapsto X_{\leq}(y, k^{\pm}(y))$ are non-decreasing, and moreover that $\int_y^{y'} d\nu > 0$ implies $X_{\leq}(y, k^+(y)) \subseteq X_{<}(y', k^-(y'))$. Then k^+ is right continuous, k^- is left continuous, and they agree throughout $\operatorname{spt}\nu$ *except perhaps at countably many points. Setting $F(x) = y$ for each $x \in X(y, k^+(y))$ defines $F : X \longrightarrow Y$ [μ-a.e.], and $\gamma = (id \times F)_{\#}\mu$ is the unique maximizer of* (1) *on $\Gamma(\mu,\nu)$.*

(c) In addition, assume $\operatorname{spt}\nu$ *is connected, say* $\operatorname{spt}\nu = [\underline{y}, \bar{y}]$*. Then F agrees μ-a.e. with the continuous function*

$$\bar{F}(x) = \begin{cases} \bar{y} \text{ if } x \notin \bigcup_{y \in (\underline{y}, \bar{y})} X_{<}(y, k^-(y)) \\ y \text{ if } x \in X_{\leq}(y, k^+(y)) \setminus X_{<}(y, k^-(y)) \text{ with } y \in (\underline{y}, \bar{y}), \\ \underline{y} \text{ if } x \in \bigcap_{y \in (\underline{y}, \bar{y})} X_{\leq}(y, k^+(y)). \end{cases} \quad (13)$$

(d) If, in addition, $Y \subset$ $\operatorname{spt}\nu$ *then $\bar{F} : X \longrightarrow Y$.*

Our strategy for proving (b) is to show the $\operatorname{spt}\gamma$ is contained in the the *s^*-subdifferential*

$$\partial^{s^*} v^* := \{(x,y) \in \overline{X} \times \overline{Y} \mid v(y') \geq v(y) - s(x,y) + s(x,y') \quad \forall y' \in \overline{Y}\}.$$

of the Lipschitz function $v : \overline{Y} \longrightarrow \mathbf{R} \cup \{+\infty\}$ solving the differential equation $v'(y) = k^+(y)$ a.e. From there, we can conclude optimality of γ using well-known results which show s^*-subdifferentials have a property (known as *s-cyclical monotonicity*) characterizing the support of optimizers [20] [51]. We also give a self-contained duality-based proof of optimality of γ as a byproduct of our uniqueness argument.

Proof. (a) Proposition 2 implies $X(y, k)$ is the intersection with X of an $m-1$ dimensional Lipschitz submanifold (orthogonal to $D_x s_y(x, y) \neq 0$ wherever the latter is defined and continuous). Since both μ and ν vanish on hypersurfaces, the function

$$h(y, k) := \mu[X_{\leq}(y, k)] - \nu[(-\infty, y)]$$

is continuous, and for each $y \in Y$ climbs monotonically from $-\nu[(-\infty, y)]$ to $1 - \nu[(-\infty, y)]$ with $k \in \mathbf{R}$. The intermediate value theorem then implies the existence of $k^{\pm}(y)$. Continuity of $h(y, k)$ also confirms that the zero set $[k^-(y), k^+(y)]$ of $k \mapsto h(y, k)$ is closed, that k^- is lower semicontinuous and k^+ is upper semicontinuous.

(b) Observe $k^-(y) < k^+(y)$ implies the open set $X_{<}(y, k^+(y)) \setminus X_{\leq}(y, k^-(y))$ is non-empty, because the image of $x \in X \mapsto s_y(x, y)$ is connected. This open subset of X is disjoint from $X_{<}(y', k^+(y')) \setminus X_{\leq}(y', k^-(y'))$ whenever $\nu[(y, y')] \neq 0$, by the monotonicity assumed of $X_{\leq}(\cdot, k^{\pm}(\cdot))$. Since X can only contain countably many disjoint open sets, we conclude $k^+(y) = k^-(y)$ for $y \in \operatorname{spt}\nu$, apart perhaps from countably many points.

Having established $\nu[\{y \mid k^+ > k^-\}] = 0$, we shall continue the proof by showing the solution

$$v(y) := \int^y k^+(y) dy.$$

of $v'(y) = k^+(y)$ has a s^*-subdifferential which is closed and contains S^+ defined by

$$S^{\pm} := \{(x, y) \in X \times Y \mid s_y(x, y) = k^{\pm}(y)\}.$$

(A similar argument shows the solution to $v' = k^-$ has a s^*-subdifferential containing S^-.) Note that the global bound $s_y \in L^\infty$ implies v is Lipschitz, hence differentiable Lebesgue almost everywhere.

To begin, suppose $(x', y') \in S^+$ and $k^+(y') = k^-(y')$, so that v is differentiable at y' and $v'(y') = s_y(x', y')$. Since $X_{\leq}(y', v'(y')) \subseteq X_{\leq}(y' + \delta, v'(y' + \delta))$ for almost all $\delta > 0$, we see $s_y(x', y) \leq v'(y)$ for almost all $y > y'$. Integrating over $y \in [y', y^+]$ yields

$$v(y^+) - s(x', y^+) \geq v(y') - s(x', y') \tag{14}$$

for all $y^+ \geq y'$. On the other hand, since $v'(y)$ is continuous at y' and $X(y, v'(y))$ is a hypersurface in the open set X, we deduce $x' = \lim_{i\to\infty} x_i$ for some sequence satisfying $s_y(x_i, y') > v'(y')$. Then from $X_{\leq}(y-\delta, v'(y-\delta)) \subset X_{\leq}(y, v'(y))$ we deduce $s_y(x_i, y'-\delta) > v'(y'-\delta)$ for almost all $\delta > 0$ hence $s_y(x', y'-\delta) \geq v'(y'-\delta)$. Integrating over $y - \delta \in [y^-, y']$ yields

$$s(x', y') - v(y') \geq s(x', y^-) - v(y^-) \tag{15}$$

for all $y^- \leq y'$. Together (14)–(15) show $(x', y') \in \partial^{s^*} v^*$.

It remains to consider the countable collection of points $y \in \operatorname{spt} \nu$ where $k^+(y) > k^-(y)$. We claim k^+ is right continuous. Since the s-subdifferential $\partial^{s^*} v^*$ is closed, this will complete the proof that $S^+ \subset \partial^{s^*} v^*$. Due to the semicontinuity already established, it is enough to show

$$k^+(y) \leq L^+ := \liminf_{\delta \downarrow 0} k^+(y + \delta).$$

For any sequence $y_i \geq y_{i+1}$ decreasing to y with $k^+(y_i) \to L^+$, the sets $X_{\leq}(y_i, k^+(y_i))$ decrease by assumption. The limit set $X_\infty := \bigcap_{i=1}^\infty X_{\leq}(y_i, k^+(y_i))$ therefore contains $X(y, k^+(y)) \subseteq X_\infty$. On the other hand, from the definition and continuity of s_y we check $X_\infty \subseteq X_{\leq}(y, L^+)$. Monotonicity of $k \mapsto X_{\leq}(y, k)$ yields $k^+(y) \leq L^+$ as desired. The proof that k^- is left continuous is similar, and implies s-cyclical monotonicity of S^-.

Since $k^\pm$ are both constant on any interval $I \subset \mathbf{R}$ with $\nu[I] = 0$, we conclude they are continuous functions except perhaps at the countably many points in $\operatorname{spt} \nu$ where they disagree (k^- being left continuous and k^+ right continuous at such points). We shall show $\gamma := (id \times F)_\# \mu$ is well-defined and maximizes (1).

The definitions of F and S^+ were chosen to ensure $Graph(F) \subset S^+$. There are two kinds of irregularities to consider. Any discontinuities in k^+ correspond to countably many gaps

$$X_<(y, k^+(y)) \setminus X_<(y, k^-(y)) \tag{16}$$

in the distribution μ of wives. Although F has not been defined on these open gaps, they have μ measure zero. In addition, $Y \setminus \operatorname{spt} \nu$ may consist of at most countably many intervals $(\underline{y}_i, \bar{y}_i)$. To each such interval corresponds a relatively closed gap

$$X_{\leq}(\bar{y}_i, k^+(\bar{y}_i) \setminus X_<(\underline{y}_i, k^-(\underline{y}_i)) \tag{17}$$

of μ measure zero on which F is not generally well-defined; typically this gap consists of a single Lipschitz submanifold, throughout which we have attempted to simultaneously assign F every value in $[\underline{y}_i, \bar{y}_i]$.

Apart from these two kinds of gaps (open and closed) — which have μ measure zero and on which F may not be (well-)defined — the $m-1$ dimensional Lipschitz submanifolds $X(y, k^+(y))$ foliate the support of μ in X. The leaves of this foliation correspond to distinct values $y \neq y'$ and are disjoint since $\nu([y, y']) > 0$ implies $X(y, k^+(y)) \subset X_<(y', k^+(y'))$ disjoint from $X(y', k^+(y'))$. This shows F is defined μ-a.e.; it is also μ-measurable since $F^{-1}((-\infty, y])$ differs from a relatively closed subset of X only by the countably many above-mentioned gaps, which are themselves Borel and have vanishing μ measure. Since F was selected so that $F_{\#}\mu := \mu \circ F^{-1}$ agrees with ν on the subalgebra of intervals $(\infty, y]$ parameterized by $y \in Y$, we see $\gamma := (id \times F)_{\#}\mu$ lies in $\Gamma(\mu, \nu)$ and is well-defined. Since $Graph(F) \subset S^+ \subset \partial^{s^*} v^*$, we conclude γ vanishes outside $\partial^{s^*} v^*$. On the other hand, s^*-subdifferentials are known to be s-cyclically monotone [43], whence optimality of γ follows from standard results [51]. We later give a self-contained proof, as a byproduct of our uniqueness argument below.

(c) Before addressing uniqueness, let us consider point (c). Each point $x \in X$ outside the μ-negligible gaps (16)–(17) belongs to a leaf $X(y, k^+(y))$ of the foliation corresponding to a unique $y \in Y$; moreover $k^+(y) = k^-(y)$. Since $X(y, k^+(y)) = X_\leq(y, k^+(y)) \setminus X_<(y, k^-(y))$ for such y we see $F = \bar{F}$ holds μ-a.e. Let us now argue $\bar{F}$ is continuous.

The fact that $\operatorname{spt} \nu = [\underline{y}, \bar{y}]$ is connected rules out all gaps (17) of the second kind in X excepting two, corresponding to the intervals in Y to the left of $\underline{y}$ and to the right of $\bar{y}$. Thus $\bar{F}$ is well-defined throughout X, taking constant values $\underline{y}$ and $\bar{y}$ on each of these two gaps, and the constant value $\bar{F}(y) = y$ on each remaining open gap (16). From (13) and the monotonicity assumed of $y \in Y \mapsto X_\leq(y, k^\pm(y))$ we see

$$\bar{F}^{-1}((-\infty, y]) = \begin{cases} \emptyset & \text{if} \quad y < \underline{y}, \\ \bigcap_{y' \in (\underline{y}, \bar{y})} X_\leq(y', k^+(y')) & \text{if} \quad y = \underline{y}, \\ X_\leq(y, k^+(y)) & \text{if}\; y \in (\underline{y}, \bar{y}), \\ X & \text{if} \quad y \geq \bar{y}, \end{cases}$$

and

$$\bar{F}^{-1}([y, \infty)) = \begin{cases} X & \text{if} \quad y \leq \underline{y}, \\ X_\geq(y, k^-(y)) & \text{if}\; y \in (\underline{y}, \bar{y}), \\ X \setminus \bigcup_{y' \in (\underline{y}, \bar{y})} X_<(y', k^-(y')) & \text{if} \quad y = \bar{y}, \\ \emptyset & \text{if} \quad y > \bar{y}, \end{cases}$$

are relatively closed, hence $\bar{F}$ is continuous.

(d) is obvious from the monotonicity of $X_\leq(y, k^+(y))$. ■

Proof of uniqueness in Theorem 4(b). To prove unicity, we show that the function v constructed above, along with

$$u(x) := \sup_{y \in Y} s(x,y) - v(y) \tag{18}$$

solve the dual problem (5). (Since v is Lipschitz, we may equivalently take the supremum defining u over Y^0 or $\overline{Y}$.) From there we proceed to argue along well-trodden lines. From (18) we see (6) holds on $X \times Y$. Moreover, the fact that $Graph(F) \subset \partial^{s^*} v^*$ shows for $x \in \mathrm{Dom}\, F$ that the supremum (18) is attained at $y = F(x)$; thus $u < \infty$ on $\mathrm{Dom}\, F$. This implies v has a global lower bound, which in turn implies a global upper bound for u. On the other hand,

$$v(y) \geq \sup_{x \in X} s(x,y) - u(x)$$

shows u is bounded below, hence belongs to L^∞. Since $y = F(x)$ yields equality in the stability constraint (6), integrating s against $\gamma = (id \times F)_{\#}\mu$ yields

$$\begin{aligned} MK^* \geq \int s d(id \times F)_{\#}\mu &= \int [u(x) + v(F(x))] d\mu(x) \\ &= \int u d\mu + \int v d\nu \\ &\geq MK_*, \end{aligned}$$

where we have concluded $v \in L^1(\nu)$ and hence $(u,v) \in Lip_s$ from the second equality. Since the opposite inequality $MK^* \leq MK_*$ is immediate from (6) and $\gamma \in \Gamma(\mu,\nu)$, we conclude that our functions (u,v) attain the infimum (5). (We also obtain direct confirmation that $MK^* = MK_*$ and $(id \times F)_{\#}\mu$ attains the supremum (1).)

Now let $\bar{\gamma}$ be any other optimizer for (1). To establish uniqueness, we shall show $\bar{\gamma}$ vanishes outside the graph of F, after which [1, Lemma 3.1] equates $\bar{\gamma}$ with $(id \times F)_{\#}\mu$ to conclude the proof. The set $S := \mathrm{spt}\, \gamma \cap (X \times \mathrm{Dom}_0\, Dv)$ carries the full mass of $\bar{\gamma}$. Each $(x', y') \in S$ produces equality in (6), hence $s_y(x', y') = v'(y')$. Thus $x' \in X(y', k(y'))$ and $F(x') = y'$, showing $S \subset Graph(F)$ as desired. ■

4.2 Universally nesting surpluses reduce dimension

Nestedness is generally a property of the three-tuple (s, μ, ν); that is, for most non-degenerate surplus functions, the model may or may not be nested depending on the measures under consideration. In some cases, however, the surplus function is such that the model is nested for all measures (μ, ν). We show this occurs precisely when the surplus takes *pseudo-index* form; for $s \in C^2$ this means there exist C^1 functions α and I on $X \subset \mathbf{R}^m$ and σ on $I(X) \times Y \subset \mathbf{R}^2$ such that

$$s(x, y) = \alpha(x) + \sigma(I(x), y). \tag{19}$$

In this case the different components of $x = (x_1, \cdots, x_m)$ are relevant only in so far as the determine the value of the index function $I(x)$, and the dimension of the transport problem is effectively reduced from $m+1$ to $1+1$, since it becomes equivalent to matching $I_{\#}\mu$ with ν optimally for σ. For connected domains, we shall also see non-degeneracy implies the effective surplus function σ is either super- or sub-modular.

Proposition 5 (Surpluses nesting universally are pseudo-index) *Assume $X \subseteq \mathbf{R}^n$ and $Y = (a, b) \subset \mathbf{R}$ are open and connected, and that the surplus $s \in C^2(X \times Y)$ is non-degenerate, meaning $D_x s_y$ is nowhere vanishing on $X \times Y$. Then s takes pseudo-index form* (19) *if and only if (s, μ, ν) is nested for every choice of absolutely continuous probability measures μ on X and ν on Y.*

Under the additional hypothesis that each level set $X(y, k)$ is connected, the "if" assertion in the preceding proposition follows immediately from Theorem 3.2 in [41] (which, under the additional connectedness assumption, asserts that if the optimal map is continuous for all marginals, with densities bounded above and below, then s must be of index form) and Theorem 4: if (s, μ, ν) is nested for each (μ, ν), Theorem 4 implies that the optimal map is continuous whenever ν has connected supported, contradicting the result in [41].

We offer a self-contained proof in an appendix. Aside from eliminating the connectedness assumption on the $X(y, k)$, our proof has the additional advantage of being elementary; the proof of Theorem 3.2 in [41] relies on a sophisticated result of Ma, Trudinger and Wang [34], which asserts that s-convexity of the target is a necessary requirement for optimal maps to be continuous for all marginals with densities bounded above and below.

5 Criteria for nestedness

Theorem 4 illustrates the powerful implications of nestedness, when it is present. In this section, we exploit the machinery of level set dynamics to derive several alternative characterizations of nestedness which — for sufficiently smooth data — may be easier to check in practice. The first of these asserts that nestedness is essentially equivalent to $X_{\leq}(y, k^+(y))$ expanding outward at each point on its boundary as y is increased, assuming its boundary hits ∂X transversally. To describe this outward expansion requires us to derive an analytic expression for the normal velocity of $X_{\leq}(y, k^+(y))$; this normal velocity also appears naturally in the integrodifferential analog of the Monge-Ampère type equation adapted to multi-to-one dimensional transport which is derived at (31) below. A second characterization, also requiring transversality, asserts nestedness is equivalent to the existence of a unique mapping $F : X \longrightarrow Y$ such that $y = F(x)$ splits the population proportionately to x for each $x \in X$.

We begin with a preparatory lemma. Anticipating our later needs, we allow for the possibility of matching between the line and a Riemannian manifold M with metric tensor g_{ij} of Sobolev regularity (such as $M = \partial X \subset \mathbf{R}^m$ with $\partial X \in C^1 \cap W^{2,1}$); until then the reader may imagine (M, g_{ij}) represents Euclidean space. Hausdorff measure $\mathcal{H}^d$ of dimension $d \leq m$ on M, functions of bounded variation, sets of finite perimeter and their reduced boundaries are defined as in, e.g. [14] for $M = \mathbf{R}^n$ and [15] for the general case.

Lemma 6 (Motion of sublevel sets) *Let M be a complete m-dimensional manifold with a Riemannian metric tensor $g_{ij} \in (C \cap W^{1,1})(M)$. Let $X \subset M$ and $Y \subset \mathbf{R}$ be domains of finite perimeter. Fix $s \in C^{0,1}$ nondenegerate throughout $\overline{X} \times \overline{Y}$, with $s_y \in C^1(\overline{X} \times \overline{Y})$ and each component of $(D_x s_y, s_{yy})$ in $C(Y; (C \cap W^{1,1})(X))$. Then*

$$N := \{(y, k) \in \overline{Y} \times \mathbf{R} \mid \mathcal{H}^{m-1}\left[\overline{X}(y, k) \cap \partial^* X\right] > 0\} \tag{20}$$

is closed and $N \cap \{y\} \times \mathbf{R}$ is countable for each $y \in \overline{Y}$, where $\partial^* X \subset \partial X$ denotes the reduced boundary of X. Setting $U := \overline{Y} \times \mathbf{R}$, the area $\mathcal{H}^{m-1}[X(y, k)]$ varies continuously on $U \setminus N$.

If $f \in L^\infty(X \times Y)$ and $f_y := \frac{\partial f}{\partial y} \in L^\infty$, then

$$\Phi(y, k) := \int_{X_{\leq}(y,k)} f(x, y) d\mathcal{H}^m(x)$$

defines a Lipschitz function on $U := \overline{Y} \times \mathbf{R}$, with $\|\Phi\|_{C^{0,1}(U)}$ controlled by the volume and perimeter of X, $\|(f, f_y, s_{yy}, |D_x s_y|^{-1})\|_{L^\infty(X \times Y)}$ and

$$\sup_{(y,k) \in U} \mathcal{H}^{m-1}[X(y, k)] < \infty. \tag{21}$$

For a.e. $(y, k) \in U$,

$$\begin{aligned} D\Phi(y, k) &:= \left(\frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial k}\right) \\ &= \int_{X(y,k)} f \frac{(-s_{yy}, 1)}{|D_x s_y|} d\mathcal{H}^{m-1} + \int_{X_{\leq}(y,k)} \left(\frac{\partial f}{\partial y}, 0\right) d\mathcal{H}^m, \end{aligned} \tag{22}$$

where $\mp s_{yy}^{(1\pm 1)/2}/|D_x s_y|$ *are the outward normal velocities of* $X_{\leq}(y,k)$ *as either* y *or* k *is varied, the other being held fixed. If, in addition* $f \in C(Y,(C\cap W^{1,1})(X))$ *then* $\Phi \in C^1(U\setminus N)$*.*

Proof. Use the approximate Heavyside step function

$$\phi_\epsilon(t) := \begin{cases} 1 & \text{if} \quad t \leq 0, \\ 1 - t/\epsilon & \text{if}\, t \in [0,\epsilon], \\ 0 & \text{if} \quad t \geq \epsilon, \end{cases}$$

to define continuous functions

$$h_\epsilon(y,k) = \int_{\partial^* X} \phi_\epsilon(s_y(x,y)-k)\,d\mathcal{H}^{m-1}(x)$$

which are monotone in both $\epsilon > 0$ and k, and approximate

$$h^+(y,k) = \mathcal{H}^{m-1}\left[\overline{X}_{\leq}(y,k)\cap\partial^* X\right]$$

pointwise from above. This shows upper semicontinuity of h^+. By symmetry,

$$h^-(y,k) = \mathcal{H}^{m-1}\left[\overline{X}_{<}(y,k)\cap\partial^* X\right]$$

is lower semicontinuous. Since

$$(h^+ - h^-)(y,k) = \mathcal{H}^{m-1}\left[\overline{X}(y,k)\cap\partial^* X\right]$$

we see $h^+ \geq h^-$, with equality holding outside of the closed set N. Since X has finite perimeter and $\overline{X}_{\leq}(y,k)$ depends monotonically on k, for each y there can only be countably many k values with $(y,k)\in N$.

To compute the instantaneous rate of displacement of $x \in X(y,k_0)$ as a function of y near y_0, with k_0 held fixed, let $x(y) = \exp_{x_0}\lambda(y)\hat{n}_0$ denote the intersection of $X(y,k_0)$ with the geodesic parallel to outer unit normal $\hat{n}_0 := D_x s_y(x_0,y_0)/|D_x s_y|$ at $x_0 \in X(y_0,k_0)$, meaning $x(y) = x_0 + \lambda(y)\hat{n}_0$ in the Euclidean case. Applying the implicit function theorem to $s_y(x(y),y) = k_0$, we see a differentiable solution λ exists near y_0 held fixed with

$$\lambda'(y_0) = \frac{-s_{yy}}{|D_x s_y|}(x_0,y_0).$$

This gives the outer normal velocity $x'(y_0) = \lambda'(y_0)\hat{n}_0$ of $\partial X_{\leq}(y,k_0)$ at x_0. The corresponding normal velocity as k is varied with y_0 held fixed is similar and even simpler to compute.

Now suppose $f \in C(Y, C\cap W^{1,1}(X))$ temporarily, with $f_y \in L^\infty$, and define

$$\Phi_\epsilon(y,k) := \int_X f(x,y)\phi_\epsilon(s_y(x,y)-k)\,d\mathcal{H}^m(x).$$

We claim the Φ_ϵ are equi-Lipschitz approximations to Φ. The dominated convergence theorem and co-area formula [14] [15] yield

$$D\Phi_\epsilon(y,k) = \int_X (f_y,0)\phi_\epsilon(s_y - k) d\mathcal{H}^m + \frac{1}{\epsilon}\int_0^\epsilon dt \int_{X(y,k+t)} f\frac{(-s_{yy},1)}{|D_x s_y|} d\mathcal{H}^{m-1}. \tag{23}$$

Letting ∇_X denote the divergence operator on X, while $\hat{n}_X$ and $\hat{n}_{X_=} := D_x s_y / |D_x s_y|$ denote the outward unit normals to X and $X_\leq(y, s_y(x,y))$ respectively, the generalized Gauss-Green formula [25, Proposition 5.8] asserts

$$\int_{X(y,k)} V \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1} = \int_{X_\leq(y,k)} \nabla_X \cdot V d\mathcal{H}^m - \int_{\partial^* X \cap \overline{X_\leq(y,k)}} V \cdot \hat{n}_X d\mathcal{H}^{m-1} \tag{24}$$

for any continuous Sobolev vector field V on $\overline{X}$. We are interested in applying this to vector fields which depend on an additional parameter $y \in Y$, whose components in local coordinates lie in $C(Y, (C \cap W^{1,1})(X))$. We claim this integral then depends continuously on $(y,k) \in U \setminus N$, and its magnitude is bounded throughout U by

$$\left| \int_{X(y,k)} V \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1} \right| \leq \|V\|_{W^{1,1}(X)} + \|V\|_{L^\infty(X)} \mathcal{H}^{m-1}(\partial^* X). \tag{25}$$

The continuous dependence on (y,k) follows from the dominated convergence theorem, and the fact that $1_{X_\leq(y',k')}$ converges to $1_{X_\leq(y,k)}$ Lebesgue almost everywhere on X and $\mathcal{H}^{m-1}$-a.e. on $\partial^* X$ as $(y',k') \to (y,k) \notin N$, in view of (20). In particular, we use this argument to show the last summand in

$$\left| \int_{X_\leq(y',k')} \nabla_X \cdot V(x,y') d\mathcal{H}^m(x) - \int_{X_\leq(y,k)} \nabla_X \cdot V(x,y) d\mathcal{H}^m(x) \right|$$
$$\leq \int_{X_\leq(y',k')} |\nabla_X \cdot (V(y') - V(y))| d\mathcal{H}^m + \int_{X_\leq(y,k) \Delta X_\leq(y',k')} |\nabla_X \cdot V(y)| d\mathcal{H}^m$$

vanishes in the limit $(y',k') \to (y,k)$, where Δ denotes the symmetric difference of the domains of integration; the other summand vanishes by the continuous dependence in $L^1(X)$ of $\nabla_X \cdot V$ on y'.

Choosing $V = \hat{n}_{X_=}$ in (24) demonstates the continuity of $\mathcal{H}^{m-1}[X(y,k)]$ on $U \setminus N$. Alternately, choosing $V = f s_{yy}^{(1\pm1)/2} D_x s_y / |D_x s_y|^2$ for fixed $y \in \overline{Y}$, we have just shown the inner integrals in (23) depend continuously on small t as long as $(y,k) \notin N$, and are locally uniformly bounded (25) with a constant depending only on $\mathcal{H}^{m-1}(\partial^* X)$ and

$$\sup_{y \in Y} \left\| \frac{f s_{yy}^{(1\pm1)/2} D_x s_y}{|D_x s_y|^2} \right\|_{(C \cap W^{1,1})(X)} < \infty. \tag{26}$$

Thus Φ_ϵ converges locally uniformly on U to a Lipschitz limit Φ_0, and $D\Phi_\epsilon$ converges pointwise on $U \setminus N$ to

$$D\Phi_0(y,k) = \int_{X_{\le}(y,k)} (\frac{\partial f}{\partial y}, 0) d\mathcal{H}^m + \int_{X(y,k)} f \frac{(-s_{yy}, 1)}{|D_x s_y|} d\mathcal{H}^{m-1}. \qquad (27)$$

Since this surface intregral is controlled by the same constants, the preceding arguments show $\Phi_0 \in C^{0,1}(U) \cap C^1(U \setminus N)$. Since N is Lebesgue negligible, $\|\Phi_0\|_{C^{0,1}(U)} = \|\Phi_0\|_{C^1(U\setminus N)}$; (27) shows the latter to be controlled by the listed constants.

On the other hand, the co-area formula [14] yields

$$\begin{aligned}|\Phi_\epsilon(y,k) - \Phi(y,k)| &\le \int_{X_{\le}(y,k+\epsilon)\setminus X_{\le}(y,k)} |f(x,y)| d\mathcal{H}^m(x) \\ &= \int_0^\epsilon dt \int_{X(y,k+t)} \frac{|f|}{|D_x s_y|} d\mathcal{H}^{m-1}(x) \\ &\le \epsilon \left\| \frac{f}{|D_x s_y|} \right\|_\infty \max_{(y,k)\in U} \mathcal{H}^{m-1}[X(y,k)]\end{aligned}$$

showing $\Phi = \Phi_0$ on U and establishing the lemma for $f \in C^{0,1}(X \times Y)$.

We handle the case $f, f_y \in L^\infty(X \times Y)$ by approximation: mollification yields a sequence $f^\delta \in C^{0,1}(X \times Y)$ with (f^δ, f^δ_y) uniformly bounded and converging to (f, f_y) pointwise a.e. as $\delta \to 0$. The dominated convergence theorem asserts pointwise convergence of

$$\Phi^\delta(y,k) := \int_{X_{\le}(y,k)} f^\delta(x,y) d\mathcal{H}^m(x)$$

to Φ. On the other hand, the version of the lemma established above shows the Φ^δ to be Lipschitz on U with a constant independent of $\delta > 0$. Thus they converge uniformly to a limit Φ sharing the same Lipschitz constant. The lemma also establishes (22) on $U \setminus N$, with (Φ^δ, f^δ) in place of (Φ, f). We would like to use the dominated convergence theorem to pass to the limit $\delta \to 0$ in (22) for a.e. (y,k). This works immediately when f is continuous. Otherwise, let $Z \subset X \times Y$ denote the Lebesgue negligible set where f^δ fails to converge to f. Fubini's theorem shows $Z(y) := \{x \in X \mid (x,y) \in Z\}$ has zero measure for a.e. $y \in Y$. Applied to its indicator function, the co-area formula

$$0 = \int_X 1_{Z(y)} d\mathcal{H}^m = \int_{\mathbf{R}} dk \int_{X(y,k)} \frac{1_{Z(y)}}{|D_x s_y|} d\mathcal{H}^{m-1}$$

then yields $\mathcal{H}^{m-1}$ negligibility of $Z(y) \cap X(y,k)$ for a.e. k. For such y and k, the dominated convergence theorem permits passage to the $\delta \to 0$ limit in (22), to complete the proof. ■

Our next goal is to establish Theorem 7, which describes how the set of wives hypothetically paired with husband $y \in Y \subset \mathbf{R}$ move in response to changes in his type.

Theorem 7 (Dependence of iso-husbands on husband type) *(a) Let $X \subset \mathbf{R}^m$ and $Y \subset \mathbf{R}$ be connected open sets of finite perimeter, equipped with Borel probability measures $d\mu(x) = f(x)dx$ and $d\nu(y) = g(y)dy$ whose Lebesgue densities satisfy $\log f \in L^\infty(X)$ and $\log g \in L^\infty_{loc}(Y)$. Assume $s \in C^2$ is non-degenerate throughout $\overline{X} \times Y$, with all components of $(D_x s_y, s_{yy})$ in $C(Y; C \cap W^{1,1}(X))$ Then the functions $k^\pm$ of Theorem 4(a) coincide. Moreover, $k := k^\pm \in C^{0,1}_{loc}(Y)$ and $|k'(y)| \leq \|s_{yy}\|_{L^\infty(X)} + g(y)\|D_x s_y/f\|_{L^\infty(X)}/\mathcal{H}^{m-1}(X(y,k(y)))$ a.e.*

(b) If, in addition $\log f \in (C \cap W^{1,1})(X)$ and $\log g \in C^0_{loc}(Y)$ then k is continuously differentiable outside the relatively closed set

$$Z := \{y \in Y \mid (y, k(y)) \in N \text{ from } (20)\} \tag{28}$$

and $k'(y) = -\frac{h_y}{h_k}(y, k(y))$ is given by (30) on $Y \setminus Z$. As $y \in Y \setminus Z$ increases the outward normal velocity of $X_\leq(y, k(y))$ at $x \in X(y, k(y))$ is given by $(k' - s_{yy})/|D_x s_y|$. If $\log g \in C(\overline{Y})$, then $k'(y)$ diverges to $+\infty$ at the endpoints of Y unless $\mathcal{H}^{m-1}[X(y, k(y))]$ remains bounded away from zero in this limit.

Proof. Since the non-degeneracy of s extends to $\overline{X} \times Y$, Proposition 2 shows $\overline{X}(y,k) := \overline{X}_\leq(y,k) \setminus \overline{X}_<(y,k)$ to be the intersection with $\overline{X}$ of an $m-1$ dimensional C^1 submanifold orthogonal to $D_x s_y(x,y) \neq 0$. As in Theorem 4, $k^\pm(y)$ represent the maximal and minimal roots of the continuous function

$$h(y,k) := \mu[X_\leq(y,k)] - \nu[(-\infty, y)], \tag{29}$$

which depends monotonically on k. The open set $X_<(y, k^+(y)) \setminus X_\leq(y, k^-(y))$ carries none of the mass of μ, hence must be empty since $\log f$ is real-valued. Thus $k^+ = k^-$ is continuous on Y and $Z \subset Y$ is relatively closed.

Under the asserted hypotheses we claim $h \in C^{0,1}_{loc}(Y \times \mathbf{R})$ and continuously differentiable outside of the set N of zero measure from (20). Indeed, h is locally Lipschitz on $Y \times \mathbf{R}$ according to Lemma 6, and its partial derivatives are given a.e.

$$h_k = \frac{\partial \Phi}{\partial k} = \int_{X(y,k)} f(x) \frac{d\mathcal{H}^{m-1}(x)}{|D_x s_y(x,y)|} \geq 0 \qquad \text{and} \tag{30}$$
$$h_y = -g + \frac{\partial \Phi}{\partial y} = -g(y) - \int_{X(y,k)} \frac{f(x) s_{yy}(x,y)}{|D_x s_y(x,y)|} d\mathcal{H}^{m-1}(x),$$

adopting the notation Φ from (22).

In case (b) these derivative are continous outside of the closed negligible set N. Since $h(y,k(y))=0$, if $h_k\neq 0$ the implicit function theorem then yields $k\in C^1_{loc}(Y\setminus Z)$, with $k'=-h_y/h_k$, and the stated bounds follow. In case (a), the Clarke implicit function theorem yields the required bound on $k\in C^{0,1}_{loc}(Y)$ [9], provided we can provide a positive lower bound for $h_k(y,k)$ a.e. in a neighbourhood of the graph $y=k(y)$ over compact subsets of Y. Since $N\subset\overline{Y}\times\mathbf{R}$ has measure zero, such a bound follows from (30) provided we obtain a positive lower bound for $\mathcal{H}^{m-1}[X(y,k)]$. But this comes from lower semicontinuity of the relative perimeter of $X_\leq(y,k)$ in X [14], given that $1_{X_\leq(y',k')}\to 1_{X_\leq(y,k)}$ Lebesgue a.e. as $(y',k')\to(y,k)$, and the fact that the $m-1$ dimensional submanifold $X(y,k(y))$ is non-empty due to the connectedness of X. If $\mathcal{H}^{m-1}[X(y,k(y)]$ tends to zero at either endpoint of Y, then h_k tends to zero but $h_y(y,k(y))\to -g(y)$ does not, showing $k'=-h_y/h_k$ diverges.

On compact subsets of Y, the outer normal velocity of $\partial X_\leq(y_0,k(y_0))$ at x_0 comes from applying Lemma 6 with $s(x,y)-\int^y k$ in place of s; (the $\|s_{yy}\|_{W^{1,1}}$ bound hypothesized in that lemma is not needed for this particular claim). ∎

Corollary 8 (Dynamic criterion for nestedness) *Adopting the hypotheses and notation of Theorem 7(b): If the model is nested then $k'-s_{yy}\geq 0$ for all $y\in Y\setminus Z$ and $x\in X(y,k(y))$, with strict inequality holding at some x for each y. Conversely, if $Z=\emptyset$ and strict inequality holds for all $y\in Y$ and $x\in X(y,k(y))$, then the model is nested.*

Proof. Away from Z, continuous differentiability of $k=k^\pm$ and the fact that the outward normal velocity of $X_\leq(y,k(y))$ at $x\in X(y,k(y))$ is given by $\frac{k'-s_{yy}}{|D_x s_y|}$ are established in Theorem 7. Differentiating $\nu[(-\infty,y)]=\mu[X_\leq(y,k(y))]$ gives

$$\begin{aligned} g(y) &= \frac{d}{dy}\int_{X_\leq(y,k(y))} f(x)d^m x \\ &= \int_{X(y,k(y))}\frac{k'(y)-s_{yy}(x,y)}{|D_x s_y|} f(x)d\mathcal{H}^{m-1}(x), \end{aligned} \tag{31}$$

at each $y\in Y\setminus Z$, according to Lemma 6. If the model is nested, so that $y\in Y\mapsto X_\leq(y,k(y))$ is increasing, this velocity must be non-negative at each boundary point. Positivity of g in the formula above shows $k'-s_{yy}$ must be positive at some boundary point.

Conversely, if this velocity is positive at each boundary point, it means $X_\leq(y,k(y))$ expands outwardly with y at each boundary point x, hence increases strictly with y over the interval Y. This confirms the model is nested. ■

Remark 9 (Boundary transversality of indifference sets) *If $y \in Z$ in (28), the indifference set of y must intersect ∂X in a set of positive area. This can only happen if the normal $D_x s_y$ to this indifference set coincides with the normal to $\partial^* X$ almost everywhere on their intersection.*

A sufficient condition for the set $Z \bigcap Y$ to be empty in (28) is therefore that the closure of any potential indifference set in X of a husband type $y \in Y$ intersects ∂X transversally. For a Lipschitz domain X this transversality implies the intersection is a Lipschitz submanifold of codimension 1 in ∂X via the Clarke implicit function theorem [9].

Remark 10 (Monge-Ampère type integrodifferential equation) *Identifying $k = v'$ and $X(y, k(y)) = F^{-1}(y)$, equation (31) should be compared to the Monge-Ampère type equation*

$$g(y) = \pm f(F^{-1}(y)) \det[D^2 v - D^2_{yy} s] / \det \left[D^2_{xy} s\right]_{x=F^{-1}(y)} \qquad (32)$$

which arises in the $n = m$ framework of Ma, Trudinger and Wang [34]. This comparison highlights the fact that positivity of the normal velocity $[v'' - s_{yy}]_{y=F(x)}$ plays a role in the multi-to-one dimensional problem analogous to strict ellipticity for (32). Indeed, Corollary 22 of the final section exploits this positivity to establish a Lipschitz bound on F and allow us to initiate our bootstrap to higher regularity. In the same section, (49) shows the Jacobian version of the balance condition (31) analogous to (4) takes the form expected from the co-area formula, namely

$$g(y) = \int_{F^{-1}(y)} \frac{f(x)}{|DF(x)|} d\mathcal{H}^{m-1}(x).$$

Corollary 11 (Unique splitting criterion for nestedness) *A model (s, μ, ν) satisfying the hypotheses of Theorem 7(b) with $Z \bigcap Y = \emptyset$ is nested if and only each $x \in X$ corresponds to a unique $y \in Y$ splitting the population proportionately, i.e. which satisfies*

$$\int_{X_{\leq}(y, s_y(x,y))} d\mu = \int_{-\infty}^{y} d\nu. \qquad (33)$$

In this case, the optimal map from μ to ν is given by $F(x) = y$.

Proof. First assume the model is nested, and suppose both (x, y) and (x', y') satisfy (33), with $y \neq y'$, so $k(y) = s_y(x, y)$ and $k(y') = s_y(x', y')$.

Taking $y < y'$ without loss of generality, the hypothesis $g > 0$ of Theorem 7 combines with the conclusion $x \in X_{\leq}(y, k(y)) \subset X_{<}(y', k(y'))$ of Theorem 4 to imply $s_y(x, y') < k(y') = s_y(x', y')$. Thus $y \neq y'$ implies $x \neq x'$ as desired.

Conversely, suppose each $x \in X$ corresponds to a single y splitting the population proportionately. To prove the model is nested, let us first check that the outward velocity $(k' - s_{yy})/|D_x s_y| \geq 0$ of $X_{\leq}(y, k(y))$ at $x \in \partial X_{\leq}(y, k(y))$ given by Theorem 7 is non-negative for all $y \in Y$; the moving boundary is a $C^{0,1}$ submanifold of X according to the non-degeneracy of Proposition 2.

To derive a contradiction, suppose $k' - s_{yy} < 0$ for some $y_0 \in Y$ and $x_0 \in X(y_0, k(y_0))$. Continuity shows this remains true for all nearby y and $x \in X(y, k(y))$, thus the boundary of $X_{\leq}(y, k(y))$ moves continuously inward near x_0 as y increases through y_0. Each point x' in a sufficiently small neighbourhood $N_r := B_r(x_0) \setminus X_{\leq}(y_0, k(y_0))$ therefore belongs to $X(y', k(y'))$ for some $y' < y_0$. On the other hand,

$$\lim_{y \uparrow \bar{y}} \mu[X_{\leq}(y, k(y))] = \lim_{y \uparrow \bar{y}} \nu[(\infty, y)] = 1$$

where $\bar{y} := \sup Y$. Since $\mu(N_r) > 0$ by hypothesis, we conclude N_r intersects $X_{\leq}(y, k(y))$ for $y > y_0$ sufficiently close to $\bar{y}$. Fix x' from this intersection and let y'' denote the infimum of points $y > y_0$ such that $x' \in X_{\leq}(y, k(y))$. Then $x' \in X(y', k(y')) \cap X(y'', k(y''))$ for $y'' \in (y_0, \bar{y})$, with both $y' < y_0$ and y'' splitting the population proportionately at x', the desired contradiction.

This establishes $y \in Y \mapsto X_{\leq}(y, k(y))$ is monotone non-decreasing. It remains to confirm this monotonicity is strict. Unless $X_{\leq}(y, k(y)) \subset X_{<}(y', k(y'))$ for each pair of points $y < y'$ in Y, some x lies in the boundary of both sets. But then both y and y' split the population proportionately at x. This contradiction implies the model is nested, and Theorem 4 implies the stable matching is given by $F(x) = y$. ■

Remark 12 (Methodological limitations are sharp) *This corollary confirms that nestedness is essentially a sharp condition for our method of solving the problem to work; in its absence the matching function F given by the procedure above fails to be well-defined: when the level sets $X(y, k)$ and $X(\bar{y}, \bar{k})$ selected to match with $y \neq \bar{y}$ intersect at some x, our construction would attempt to simultaneously assign both $F(x) := y$ and $F(x) := \bar{y}$.*

6 Examples

In the next sections, we address the higher regularity of optimal maps and potentials in nested problems. Before doing so, we discuss several related examples demonstrating the phenomena we subsequently analyze. The first is nested, the second is not and the third explores the transition. All involve maximizing the bilinear surplus $s(x,y) = x \cdot y$ between two probability measures on $\mathbf{R}^m$, which is equivalent to minimizing the quadratic cost $c(x,y) = \frac{1}{2}|x-y|^2$. In the first example, the target measure will be supported on a segment; in the second and third, it will be supported on a circular arc. These examples admit unique solutions given explicitly by optimal maps, but demonstrate that such maps and the corresponding potentials u and v will not necessarily be smooth at the boundary. Although these examples are solved using classical methods and symmetry, they guide our intuition for what to expect from problems which do not admit explicit solution.

Besides arising frequently in different applications [12] [36] [44] [17] [24], the bilinear surplus / quadratic cost is the easiest objective to analyze when $mn > 1$. For these reasons it has played a seminal role in theoretical developments [28] [4] [5] [6]. With this surplus function, target measures supported on lower dimensional sets were considered implicitly in [10] [45] [35] and explicitly in [21]. The basic result dating to these early works is that the payoff functions u, v which optimize (5) may be taken to be convex on $\mathbf{R}^m$, and that γ optimizes (1) if and only if it is supported in the subdifferential

$$\partial u := \{(x,y) \in \mathbf{R}^m \times \mathbf{R}^m \mid u(z) \geq u(x) + y \cdot (z - x) \quad \forall z \in \mathbf{R}^m\}$$

of some convex $u : \mathbf{R}^m \longrightarrow \mathbf{R} \cup \{+\infty\}$. In particular, the optimal γ is unique provided μ vanishes on Lipschitz hypersurfaces [35], or at least on those hypersurfaces parameterized by convex differences [20] [22].

Example 13 (From convex volumes to segments) *Fix $s(x,y) = x \cdot y$ on $\mathbf{R}^m \times \mathbf{R}^m$, and consider transporting volume from a convex body $X \subset \mathbf{R}^m$ to a uniform meassure on a subsegment of the x_1-axis. In this case the convex payoff function $u(x_1, \ldots, x_m) = U(x_1)$ turns out to depend on x_1 only, and the iso-husband sets consist of (hyper-)planes of constant x_1; convexity ensures that — apart from the two supporting hyperplanes — these hit $\partial^* X$ transversally, so Z is empty. Since these hyperplanes do not intersect each other, Corollary 11 shows this problem is nested. The optimal map $y = F(x) = (U'(x_1), 0, \ldots, 0)$ depends monotonically on x_1, and can be found by solving a problem in single variable calculus, analogous to* (3)*. Taking X to be a ball leads to elliptic integrals not explicitly soluble, but transporting the solid paraboloid*

$$X := \{x \in \mathbf{R}^m \mid \frac{1}{2}\sum_{i=2}^{m} x_i^2 < x_1 < const\}$$

to a segment $Y = (0, L\hat{e}_1)$ of the appropriate length yields the optimal map $F(x_1, \ldots, x_m) = (x_1^{(m+1)/2}, 0, \ldots, 0)$ and potentials $u(x) = \frac{2}{m+3} x_1^{1+(m+1)/2}$ and $v(y) = \frac{m+1}{m+3} y^{1+2/(m+1)}$ explicitly. Their behaviour near the origin shows we cannot generally expect v to be better than $C^{1,\frac{2}{m+1}}$; similarly, we expect $F \in C^{\frac{m}{2},\frac{1}{2}}$ and $u \in C^{\frac{m}{2}+1,\frac{1}{2}}$ to be sharp Hölder exponents near the origin, at least for m even and sufficiently convex domains X. This lack of C^2 smoothness of v at the boundary of Y was predicted by Theorem 7; like the lack of higher order smoothness of F and u at the boundary of X, it directly reflect the unequal dimensions of the source and target.

In fact, in the absence of strong convexity of X, we cannot even expect this much smoothness up to the boundary. If we consider instead

$$X := \{x \in \mathbf{R}^m \mid \frac{1}{2}\Big(\sum_{i=2}^{m} x_i^2\Big)^k < x_1 < const\}$$

for $k \geq 1$, the optimal map takes the form $F(x) = Cx_1^{1+\frac{m-1}{2k}}$. By choosing k large enough, this shows that we cannot generally expect $F \in C^{1,\alpha}$ up to the boundary for any $\alpha > 0$, unless we assume X has some uniform convexity.

Note that the surplus in the preceding example is of pseudo-index form, with index function $I(x_1, ..., x_m) = x_1$. Generally speaking, the bilinear surplus $s(x, y) = x \cdot y$ on an open set $X \subseteq \mathbf{R}^m$ and a smooth curve $Y \subset \mathbf{R}^m$ is of index form if and only if Y is contained in a line. The examples below treat the case where Y is a circular arc and therefore s is not of pseudo-index form. Special cases of these examples were studied in [41], before the notion of nestedness had been formulated.

Example 14 (From punctured ball to punctured circle) *Let $s(x, y) = x \cdot y$ on $\mathbf{R}^m \times \mathbf{R}^m$ and consider transporting volume $\mu = \frac{1}{\mathcal{H}^m[X]}\mathcal{H}^m$ from the punctured ball $X := \{x \in \mathbf{R}^m \mid 0 < |x| < 1\}$ to arclength $\nu = \frac{1}{\mathcal{H}^1[Y]}\mathcal{H}^1$ on the punctured circle $Y = \{y \in \mathbf{R}^m \mid y_1^2 + y_2^2 = 1, \quad y_3 = \cdots = y_m = 0 < y_1 + 1\}$. Since the map $F(x) = \frac{(x_1,x_2)}{(x_1^2+x_2^2)^{1/2}}$ pushes μ forward to ν, and its graph lies in the subdifferential of the convex function $u(x) = \sqrt{x_1^2 + x_2^2}$, we see F is optimal. Moreover u and its Legendre transform*

$$v(y) = \begin{cases} 0 & \text{if } |x| \leq 1 \text{ and } x_3 = \cdots = x_m = 0, \\ +\infty & \text{else,} \end{cases}$$

minimize the dual problem (5). Although v is smooth on Y, note F and Du fail to be smooth on the codimension 2 submanifold $x_1 = 0 = x_2$, *where they are discontinuous. The iso-husband sets $F^{-1}(y)$ consist of connected hypersurfaces in X bounded by this singular set; they intersect the boundary of X transversally. On the other hand, parameterizing the punctured circle Y using $-\pi < \theta < \pi$, the surplus function takes the form $s(x,\theta) = x_1 \cos\theta + x_2 \sin\theta$, and the payoff $v(\theta) = 0$. The potential indifference sets $X(\theta,k) := \{x \in \overline{X} \mid x_1 \sin\theta - x_2 \cos\theta = k\}$ are connected. The iso-husband set $F^{-1}(\theta)$ occupies precisely half of $X(\theta, v'(\theta))$ which shows the model is not nested.*

This example illustrates a phenomenon which our subsequent analysis shows to be intimately connected with nestedness: smoothness often holds for the payoff on the lower dimensional space Y even when it fails for the optimal map and payoff on the higher dimensional space X.

For the ball X, Examples 13 and 14 represent limits of a continuum of examples consisting of circular arcs Y of angle $|\theta| < \theta_0$ and radius $1/\theta_0$. Having less symmetry, they are not explicitly solvable, but it is natural to expect they remain nested for θ_0 less than some critical value $\theta_c \in (0,\pi)$, and become un-nested otherwise. For $\theta_0 > \theta_c$ it is not at all obvious how the singularities in F may be located (though they are characterized implicitly through duality). Note that the analysis of Example 14 extends equally well to the case where X is a ball or spherical shell $\{0 < r < |x| < R\}$ instead of a punctured ball; in the latter case F will be smooth.

One last example illustrates the transition from nestedness to non-nestedness. In this example we also see the continuity of the map $F : X \longrightarrow Y$ shown in Theorem 4 need not extend to the closed set $\overline{X}$.

Example 15 (From pie slice to circular arc) *Fix $0 < \theta_0 < \pi$ and $r_0 > 0$. Let μ be uniform on the pie-shaped region $X \subset \mathbf{R}^2$ described by $r < r_0$ and $|\theta| < \theta_0$ in polar coordinates. Let ν be uniform over the circular arc $r = 1$ and $|\theta| < \theta_0$. The optimal map F and potentials u and v are the same as in Example* (14)*, by the same arguments (or by restriction). However this model is nested if and only if $\theta_0 \leq \pi/2$; the* if *implication is shown using Corollary 11, and the* only if *by the logic of Example 14.*

7 Regularity of husband's payoff

Regularity of the map F and payoff functions u and v is a notoriously delicate question which has recieved considerable attention in case $m = n > 1$, reviewed in [13] and [37]. Priori to the work of Ma, Trudinger and Wang [34], Villani [50] had described it as *"Without any doubt, the main open problem in the field"*. When $m > n$, very little is known [41].

For $m > n = 1$, Theorem 4 shows nestedness implies continuity of F on the domain interior $X = X^0$; Theorem 7 combines with Remark 9 to give conditions guaranteeing $v \in C^2_{loc}(Y^0)$; Example 13 shows we cannot expect $k = dv/dy$ to have a Hölder exponent larger than $\frac{2}{m+1}$ at the endpoints of Y.

The present section is devoted to the following theorem, which shows that higher regularity of the husband's payoff v can be inferred from that of (s, μ, ν) and ∂X under the same boundedness and transversality conditions required by Theorem 7 and Remark 9. The strategy of the proof is to iteratively combine our hypotheses with geometric measure theoretic level-set techniques to deduce sufficient smoothness of the function $h(y, k)$ from (29) whose level sets implicitly define $k = dv/dy$. The result then follows from the implicit function theorem as in Theorem 7. We exercise care to localize the hypotheses in (x, y) where possible. We discuss smoothness of the map F and wives' payoffs u in a subsequent section.

Theorem 16 (Regularity of the husband's payoff) *Fix an integer $r \geq 1$. Under the hypotheses of Theorems 4(b) and 7(b), suppose there is an interval $Y' = (y_0, y_1) \subset Y$ such that $X' \cap \partial X \in C^1$ intersects $\overline{X(y, k(y))}$ transversally for all $y \in \overline{Y'}$, where $X' = \bigcup_{y \in Y'} \overline{X(y, k(y))}$. Then $\|k\|_{C^{r,1}(Y')}$ is controlled by the following quantities, all assumed positive and finite: $\|\log f\|_{C^{r-1,1}(X')}$, $\|\log g\|_{C^{r-1,1}(Y')}$, $\|s_y\|_{C^{r,1}(X' \times Y')}$, $\|\hat{n}_X\|_{(C^{r-2,1} \cap W^{1,1})(X' \cap \partial X)}$, $\mathcal{H}^{m-1}[\partial^* X]$,*

$$\inf_{y \in Y'} \mathcal{H}^{m-1}\left[X(y, k(y))\right] \quad \text{(proximity to ends of } Y\text{)}, \tag{34}$$

$$\inf_{x \in X', y \in Y'} |D_x s_y(x, y)| \quad \text{(non-degeneracy)}, \tag{35}$$

$$\inf_{x \in X' \cap \partial X, y \in Y'} 1 - (\hat{n}_X \cdot \hat{n}_{X_=})^2 \quad \text{(transversality)} \tag{36}$$

where $\hat{n}_{X_=}(x, y) = D_x s_y / |D_x s_y|$, and $\mathcal{H}^{m-2}\left[\overline{X(y_0, k(y_0))} \cap \partial X\right]$.

To establish this theorem, we study the motion of the interior $X_\leq(y, k)$ and boundary $X_\leq(y, k) \cap \partial X$ level sets of s_y with respect to changes in y and in k. The normal velocities for this motion are given by

$$V^\pm(x, y) := \mp \frac{(s_{yy})^{(1\pm1)/2}}{|D_x s_y|} \hat{n}_{X_=}, \qquad \hat{n}_{X_=}(x, y) := \frac{D_x s_y}{|D_x s_y|}, \tag{37}$$

and

$$V^\pm_{\partial X}(x, y) := \frac{V^\pm \cdot \hat{n}_{X_=}}{\sqrt{1 - (\hat{n}_X \cdot \hat{n}_{X_=})^2}} \hat{n}_\partial, \quad \hat{n}_\partial(x, y) := \frac{\hat{n}_{X_=} - (\hat{n}_{X_=} \cdot \hat{n}_X)\hat{n}_X}{\sqrt{1 - (\hat{n}_X \cdot \hat{n}_{X_=})^2}}, \tag{38}$$

respectively. We shall also employ the divergence operators ∇_X on X and $\nabla_{\partial X}$ on ∂X.

We remark the outward normal velocity of $X_\leq \cap \partial X$ is given by $V^\pm_{\partial X}$, since $\hat{n}_\partial$ is the outward unit normal to $X_\leq \cap \partial X$ in ∂X, and $V^\pm$ coincides with the projection of $V^\pm_{\partial X}$ onto $\hat{n}_{X_=}$. As it will play an important role in what follows, let us also remark on the smoothness of the domain X and its boundary divergence operator. Any $\partial X \in C^1$ can be parameterized locally as a graph of a function $w \in C^1(\mathbf{R}^{m-1})$; in this parameterization its metric tensor takes the form $g = I + Dw \otimes Dw$, and $\nabla_{\partial X} \cdot W = \partial_i W^i + \Gamma^i_{i\ell} W^\ell$, where the Christoffel symbol $\Gamma^i_{i\ell}$ involves first derivatives of g. Thus we need $\hat{n}_X \in W^{1,1}(\partial X)$ to define $\nabla_{\partial X}$, and if $\partial X \in C^{r-1,1}$ then $\nabla_{\partial X} \cdot W \in C^{r-3,1}$ provided the vector field W is smooth enough (say $W \in C^{r-2,1}$), with $C^{-1,1} = L^\infty$ conventionally.

Our first lemma shows the size of the indifference sets $X(y, k(y))$ and their boundaries are controlled uniformly on (y_0, y_1) by the constants listed in the theorem.

Lemma 17 (Size of indifference sets) *Under the hypotheses of Theorem 7(a), the area $A(y) = \mathcal{H}^{m-1}[X(y, k(y))]$ of the indifference set is controlled by*

$$A(y) \leq \frac{\|s_y\|_{W^{1,1}(X)}}{\inf\limits_{x \in X} |D_x s_y(x, y)|} + \mathcal{H}^{m-1}[\partial^* X]. \tag{39}$$

Under the hypotheses of Theorem 16, it satisfies

$$\sup_{y \in Y'} A(y) \leq \frac{\|D_x s_y\|_{C^{0,1}(X' \times Y')}}{\inf\limits_{(x,y) \in X' \times Y'} |D_x s_y|} \mathcal{H}^m(X') + \mathcal{H}^{m-1}(X' \cap \partial X) + \inf_{y \in Y'} A(y) \tag{40}$$

while the area of its boundary $B(y) := \mathcal{H}^{m-2}\left[\overline{X(y, k(y))} \cap \partial X\right]$ satisfies

$$\sup_{y \in Y'} B(y) \leq B(y_0) + \|\nabla_{\partial X} \cdot \hat{n}_\partial\|_{L^1(X' \cap \partial X)}, \tag{41}$$

hence both are controlled by the constants named in that theorem.

Proof. Let $X^y_\leq := \overline{X_\leq(y, k(y))}$ and $X^y_= := \overline{X(y, k(y))}$. We claim bounds on

$$A(y) := \int_{X^y_=} (\hat{n}_{X_=}(x, y) \cdot \hat{n}_{X_=}(x, y)) d\mathcal{H}^{m-1}(x) \qquad \text{and} \tag{42}$$

$$B(y) := \int_{X^y_= \cap \partial X} (\hat{n}_\partial(x, y) \cdot \hat{n}_\partial(x, y)) d\mathcal{H}^{m-2}(x). \tag{43}$$

Note both depend continuously on $y \in Y$ by Lemma 6 and Theorem 7.

The generalized Gauss-Green theorem (24) yields

$$A(y) = \int_{X^y_{\leq}} (\nabla_X \cdot \hat{n}_{X_=}) d\mathcal{H}^m - \int_{X^y_{\leq} \cap \partial^* X} (\hat{n}_{X_=} \cdot \hat{n}_X) d\mathcal{H}^{m-1}$$

from which (39) follows immediately. Now suppose the infimum (34) is attained at $z \in \overline{Y'}$. If $z \leq y$ the Gauss-Green theorem yields

$$\begin{aligned} A(y) = &\int_{X^y_{\leq} \setminus X^z_{\leq}} (\nabla_X \cdot \hat{n}_{X_=}) d\mathcal{H}^m - \int_{[X^y_{\leq} \setminus X^z_{\leq}] \cap \partial X} (\hat{n}_{X_=} \cdot \hat{n}_X) d\mathcal{H}^{m-1} \\ &+ \int_{X^z_{=}} \hat{n}_{X_=}(x,y) \cdot \hat{n}_{X_=}(x,z) d\mathcal{H}^{m-1}(x), \end{aligned}$$

where the nestedness $X^z_{\leq} \subset X^y_{\leq}$ assumed in Theorem 16 has been used. If $z > y$ a similar formula holds, with the roles of y and z interchanged. In either case (40) follows.

Similarly, the Riemannian version (24) of generalized Gauss-Green theorem asserts

$$B(y) = \int_{[X^y_{\leq} \setminus X^{y_0}_{\leq}] \cap \partial X} (\nabla_{\partial X} \cdot \hat{n}_\partial) d\mathcal{H}^{m-1} + \int_{X^{y_0}_{=} \cap \partial X} \hat{n}_\partial(x,y) \cdot \hat{n}_\partial(x,y_0) d\mathcal{H}^{m-2},$$

which implies (41). It is easy to see the quantities appearing in (40)–(41) are controlled by those listed in Theorem 16. ■

Remark 18 *If $\hat{n}_X \in C^{0,1}(\partial X)$, a similar argument using a Kirszbraun extension to ∂X can be used to bound $B(y)$ independently of $B(y_0)$.*

We next adapt Lemma 6 to the differentiation of boundary fluxes.

Lemma 19 (Derivative of a flux through a moving boundary) *Suppose $X' \subset X \subset \mathbf{R}^m$, $Y' \subset Y \subset \mathbf{R}$, (s, f, g) and k satisfy the hypotheses of Theorem 16 for some $r \geq 1$. Choose a neighbourhood $U \subset Y \times \mathbf{R}$ such that $X(y,k) \subset X'$ for all $(y,k) \in U$. If $a : X' \times Y' \longrightarrow \mathbf{R}^m$ is Lipschitz, then*

$$\Phi(y,k) := \int_{X(y,k)} a(x,y) \cdot \hat{n}_{X_=}(x,y) d\mathcal{H}^{m-1}(x) \tag{44}$$

is Lipschitz on U, with partial derivatives $(\Phi^+, \Phi^-) := (\frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial k})$ given a.e. by

$$\begin{aligned} \Phi^\pm(y,k) = &\int_{X(y,k)} [(\nabla_X \cdot a) V^\pm + \frac{1 \pm 1}{2} \frac{\partial a}{\partial y}] \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1} \\ &- \int_{\overline{X(x,y)} \cap \partial X} (a \cdot \hat{n}_X)(V^\pm_{\partial X} \cdot \hat{n}_\partial) d\mathcal{H}^{m-2}. \end{aligned} \tag{45}$$

Similarly, if $r \geq 2$ and $b : (X' \cap \partial X) \times Y' \longrightarrow T\partial X$ denotes a jointly Lipschitz family of sections of the tangent bundle of $X' \cap \partial X$, then

$$\Psi(y,k) := \int_{X(y,k)\cap\partial X} b(x,y)\cdot \hat{n}_\partial(x,y) d\mathcal{H}^{m-2}(x) \tag{46}$$

is Lipschitz on U, with partial derivatives $(\Psi^+, \Psi^-) := (\frac{\partial\Psi}{\partial y}, \frac{\partial\Psi}{\partial k})$ given a.e. by

$$\Psi^\pm(y,k) = \int_{\overline{X(y,k)}\cap\partial X} [(\nabla_{\partial X}\cdot b)V^\pm_{\partial X} + \frac{1\pm 1}{2}\frac{\partial b}{\partial y}]\cdot \hat{n}_\partial \; d\mathcal{H}^{m-2}. \tag{47}$$

Proof. To see that the proof of (47) is completely analogous to the proof of (45), let us argue the latter for $\mathbf{R}^m$ replaced by a complete Riemannian manifold M with $g_{ij} \in C \cap W^{1,1}$, such as $M = \partial X$. Note $\partial X \in C^{1,1}$ if $r \geq 2$; its lack of boundary accounts for the comparative simplicity of (47) relative to (45).

Choose a $C^{1,1}$ smooth family $a_\epsilon : X' \times Y' \longrightarrow TX'$ of sections of the tangent bundle converging to a in Lipschitz norm, so that both $\nabla_X \cdot a_\epsilon$ and $\frac{\partial}{\partial y}(\nabla_X \cdot a_\epsilon) \in L^\infty$. Define Φ_ϵ by (44) with a replaced by a_ϵ. Let $Y' = (y_0, y_1)$, so that $X \cap \partial X' = X(y_0, k(y_0)) \cup X(y_1, k(y_1))$. The generalized Gauss-Green theorem (24) theorem yields

$$\begin{aligned}\Phi_\epsilon(y,k) = &\int_{X_\leq(y,k)\setminus X_\leq(y_0,k(y_0))} \nabla_X\cdot a_\epsilon d\mathcal{H}^m \\ &- \int_{[X_\leq(y,k)\setminus X_\leq(y_0,k(y_0))]\cap\partial X} a_\epsilon\cdot\hat{n}_X d\mathcal{H}^{m-1} \\ &+ \int_{X(y_0,k(y_0))} a_\epsilon(x,y)\cdot\hat{n}_{X_=}(x,y_0)d\mathcal{H}^{m-1}(x)\end{aligned}$$

Lemma 6 then asserts Φ_ϵ is Lipschitz on U, with

$$\begin{aligned}\Phi^+_\epsilon = &\int_{X(y,k)} (\nabla_X\cdot a_\epsilon)V^+\cdot\hat{n}_{X_=} d\mathcal{H}^{m-1} + \int_{X_\leq(y,k)\setminus X_\leq(y_0,k(y_0))} \nabla_X\cdot\frac{\partial a_\epsilon}{\partial y} d\mathcal{H}^m \\ &- \int_{\overline{X(x,y)}\cap\partial X} (a_\epsilon\cdot\hat{n}_X)(V^+_{\partial X}\cdot\hat{n}_\partial)d\mathcal{H}^{m-2} - \int_{[X_\leq(y,k)\setminus X_\leq(y_0,k(y_0))]\cap\partial X} \hat{n}_X\cdot\frac{\partial a_\epsilon}{\partial y} d\mathcal{H}^m \\ &+ \int_{X(y_0,k(y_0))} \frac{\partial a_\epsilon}{\partial y}(x,y)\cdot\hat{n}_{X_=}(x,y_0)d\mathcal{H}^{m-1}(x) \\ = &\int_{X(y,k)} [(\nabla_X\cdot a_\epsilon)V^+ + \frac{\partial a_\epsilon}{\partial y}]\cdot\hat{n}_{X_=} d\mathcal{H}^{m-1} - \int_{\overline{X(x,y)}\cap\partial X} (a_\epsilon\cdot\hat{n}_X)(V^+_{\partial X}\cdot\hat{n}_\partial)d\mathcal{H}^{m-2}\end{aligned}$$

a.e.; the generalized Gauss-Green identity has again been used. A similar formula holds for Φ^-_ϵ with $(V^+, V^+_{\partial X}, \frac{\partial}{\partial y})$ replaced by $(V^-, V^-_{\partial X}, \frac{\partial}{\partial k})$. In particular, $\|\Phi_\epsilon\|_{C^{0,1}(Y\times\mathbf{R})}$ can be bounded independently of ϵ by $\|w\|_{C^{0,1}}$, $\|(V^\pm, V^\pm_{\partial X})\|_{C^0}$, (21) and (41). Since $\|a_\epsilon - a\|_{C^{0,1}(X\times Y;TX)} \to 0$ we can pass to the $\epsilon \to 0$ limit to bound $\|\Phi\|_{C^{0,1}(Y\times\mathbf{R})}$ using the same quantities, and obtain (45) from the dominated convergence theorem.

When $r \geq 2$, approximating b by b_ϵ analogously, the same reasoning yields Ψ_ϵ Lipschitz on U with

$$\Psi^\pm_\epsilon(y,k) = \int_{\overline{X(y,k)}\cap\partial X} [(\nabla_{\partial X}\cdot b_\epsilon)V^\pm_{\partial X} + \frac{1\pm 1}{2}\frac{\partial b_\epsilon}{\partial y}]\cdot \hat{n}_\partial \, d\mathcal{H}^{m-2}$$

a.e. Since $\|b_\epsilon - b\|_{C^{0,1}([X'\cap\partial X]\times Y';T\partial X)} \to 0$, we see $\|\Psi_\epsilon\|_{C^{0,1}(U)}$ is controlled as before by $\|z\|_{C^{0,1}}$, $\|V^\pm_{\partial X}\|_{C^0}$ and (41), permitting passage to the $\epsilon \to 0$ limit. ■

Corollary 20 (Iterated differentiation of fluxes) *Under the hypotheses and notation of Lemma 19, if $r \geq 2$ or $b = 0$ then*

$$(\Phi + \Psi)^\pm = \int_{X(y,k)} a^\pm \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1} + \int_{\overline{X(y,k)}\cap\partial X} b^\pm \cdot \hat{n}_\partial d\mathcal{H}^{m-2}$$

a.e. on U, where $\begin{pmatrix} a^\pm \\ b^\pm \end{pmatrix} = A_\pm \begin{pmatrix} a \\ b \end{pmatrix}$ are given by the first-order differential operators

$$A_\pm \begin{pmatrix} a \\ b \end{pmatrix} := \begin{pmatrix} (\nabla_X \cdot a)V^\pm + \frac{1\pm1}{2}\frac{\partial a}{\partial y} \\ (\nabla_{\partial X}\cdot b - \hat{n}_X \cdot a)V^\pm_{\partial X} + \frac{1\pm1}{2}\frac{db}{dy} \end{pmatrix}.$$

For each integer $0 \leq i \leq r-2$, the operator $A_\pm : B_i \longrightarrow B_{i-1}$ gives a bounded linear transformation from the Banach space

$$B_i := C^{i,1}(X' \times Y'; \mathbf{R}^m) \oplus C^{i,1}([X'\cap\partial X]\times Y'; T\partial X)$$

of $C^{i,1}$ families of sections of the tangent bundle to B_{i-1}; its norm

$$\|A_\pm\|_{B_i\to B_{i-1}} \leq C_{m,i}(1+\|V^\pm\|_{C^{i-1,1}(X'\times Y')}+(1+\|\hat{n}_X\|_{C^{i-1,1}(X'\cap\partial X)})\|V^\pm_{\partial X})\|_{C^{i-1,1}([X'\cap\partial X]\times Y')})$$

is controlled by $\|s_y\|_{C^{i,1}(X\times Y')}$, $\|\hat{n}_X\|_{C^{i-1,1}(\partial X)}$, (35) and (36). Similarly,

$$\|A_\pm\|_{C^{r-1,1}(X\times Y')\oplus\{0\}\to B_{r-2}} \leq C_{m,r-1}(1+\|V^\pm\|_{C^{r-2,1}(X\times Y')}+\|V^\pm_{\partial X}\otimes\hat{n}_X\|_{C^{r-2,1}(\partial X\times Y')})$$

is controlled by $\|s_y\|_{C^{r-1,1}(X\times Y')}$, $\|\hat{n}_X\|_{C^{r-2,1}(\partial X)}$ and (35).

Proof. The corollary follows directly from Lemma 19. The norms of the first order differential operators $A_\pm$ are elementary to estimate. ■

Finally, we are ready to iterate derivatives using the preceding lemma and its corollary to deduce Theorem 16 from Theorem 7.

Proof of theorem. Theorem 7 asserts $h(y,k) := \mu[X_{\le}(y,k)] - \nu[(-\infty, y)]$ is continuously differentiable on the set U of Lemma 19. Apart from the additive term $g(y)$, its partial derivatives $(h_+, h_-) := (g + \frac{\partial h}{\partial y}, \frac{\partial h}{\partial k})$ are given in (30) as flux integrals

$$h_\pm(y,k) = \int_{X(y,k)} fV^\pm \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1}.$$

over the indifference set $X(y,k)$. Taking $a = fV^\pm \in C^{r-1,1}(X' \times Y'; \mathbf{R}^m)$ and $b = 0$, Corollary 20 allows us to compute and bound r derivatives of $h_\pm$ on U iteratively.

Taking $0 \le i \le j \le r$, in the notation of that corollary,

$$\frac{\partial^j h_\pm}{\partial k^{j-i} \partial y^i} = \int_{X(y,k)} a_i^j \cdot \hat{n}_{X_=} d\mathcal{H}^{m-1} + \int_{\overline{X(y,k)} \cap \partial X} b_i^j \cdot \hat{n}_\partial \, d\mathcal{H}^{m-2} \qquad (48)$$

where $\begin{pmatrix} a_i^j \\ b_i^j \end{pmatrix} := A_-^{j-i} A_+^i \begin{pmatrix} fV^\pm \\ 0 \end{pmatrix} \in B_{r-j-1}$. Furthermore, $\| \begin{pmatrix} a_i^j \\ b_i^j \end{pmatrix} \|_{C^{r-1-j,1}}$ is controlled by quantities named in the theorem: $\|f\|_{C^{r-1,1}(X')}$, $\|s_y\|_{C^{r,1}(X' \times Y')}$, $\|\hat{n}_X\|_{C^{r-2,1}(X' \cap \partial X)}$, (34)–(36), $\mathcal{H}^{m-1}[\partial^* X]$ and $\mathcal{H}^{m-2}[\overline{X(y_0, k(y_0)} \cap \partial X]$, in view of Lemma 17. Since all but the final derivatives provided by the Corollary are continuous we may order them as convenient, and discover $\|h\|_{C^{r,1}(U)}$ is controlled by the listed quantities.

Now k solves $h(y, k(y)) = 0$, so the implicit function theorem result provided by Theorem 7 shows k inherits the same smoothness as h. Using (30) to bound $h_k(y, k(y)) > 0$ away from zero by the product of $\|D_x s_y / f\|_{L^\infty}$ with (34), we deduce $\|k\|_{C^{r,1}(Y')}$ is controlled by the quantities named in the theorem. ■

8 Regularity of optimal maps and potentials

Having found conditions guaranteeing smoothness of the husband's payoff $v(y)$ in the preceding section, we now turn to the wife's payoff $u(x)$ and the optimal correspondence $F : X \longrightarrow Y$ between wives and husbands.

Our main conclusions are as follows: Lipschitz continuity of F is equivalent to a strong form of nestedness, which requires a lower bound on the local speed of the motion of $F^{-1}(y)$ with respect to changes in $y \in Y$. On regions $F^{-1}([a,b])$ where this holds, higher regularity of F and u (up to the boundary) are inherited from interior regularity of v via the first order conditions

$$(Du(x), Dv(F(x))) = (D_x s(x, F(x)), D_y s(x, F(x))$$

from (7).

We begin with a logically independent proposition, which will then be combined with results of the preceding section to harvest the desired results.

Proposition 21 (Optimal maps have locally bounded variation) *The hypotheses of Theorem 4(c) and 7(a) imply $u \in C^1(X)$, $F \in (BV_{loc} \cap C)(X)$, $D_x s_y(\cdot, F(\cdot)) \in (BV_{loc} \cap C)(X)$, $F \in \mathrm{Dom}\, Dk$ on a set of $|DF|$ full measure, and*

$$(k'(F(x)) - s_{yy}(x, F(x)))DF(x) = D_x s_y(x, F(x)). \tag{49}$$

Proof. Since $k^+ = k^-$ and $\log f$ and $\log g$ are bounded on compact subsets of X and Y, we conclude F is defined and coincides with the continuous function $\bar{F}$ throughout X from Theorem 4(c).

It is well-known that the dual problem (5) admits minimizers (u, v) which inherit Lipschitz and semiconvexity bounds from $s \in C^2(X \times Y)$ [50]. Then $u \in C^{0,1}(X)$ is differentiable Lesbesgue a.e. on X, and

$$Du(x) = D_x s(x, F(x)) \tag{50}$$

for each $x \in \mathrm{Dom}\, Du$ according to (7). The right hand side is continuous and bounded, whence $u \in C^1(X)$.

Since u is also semiconvex, its directional derivatives lie in $BV(X)$. We next use (50) to deduce $F \in BV_{loc}(X)$, which means its directional weak derivatives are signed Radon measures on X. Fix $x' \in X$ and set $y' = F(x') \in Y$. Since $D_x s_y(x', y') \neq 0$, at least one of its components — say $\frac{\partial^2 s}{\partial x_1 \partial y}(x, y)$ — is non-vanishing in a neighbourhood of (x', y'). The inverse function theorem guarantees the map $(x, y) \mapsto (x, \frac{\partial s}{\partial x_1})$ has a continuously differentiable inverse defined on a neighbourhood of $(x', \frac{\partial s}{\partial x_1}(x', y'))$. From (50) and the continuity of F and Du we therefore deduce

$$F(x) = \left[\frac{\partial s}{\partial x_1}(x, \cdot)\right]^{-1} \left(\frac{\partial u}{\partial x_1}(x)\right)$$

expresses F as the composition of a C^1 map and a BV map near x'. This shows $F \in BV_{loc}(X)$ [2].

On the other hand, $k = \frac{dv}{dy} \in C^{0,1}_{loc}(Y)$ is locally Lipschitz according to Theorem 7(a). According to [2], $F \in \mathrm{Dom}\, Dk$ on a set of $|DF|$ full measure, and differentiating $k(F(x)) = s_y(x, F(x))$ yields (49) in the sense of measures; DF has no jump part since F is continuous. ■

Corollary 22 (Maps are C^1 where level set speed is non-zero) *The hypotheses of Proposition 21 imply $\|F\|_{C^{0,1}(X')} \leq \ell^{-1} \|D_x s_y\|_{C(X' \times F(X'))} < +\infty$ for any open set $X' \subset X$ having a speed limit*

$$\ell := \inf_{x \in X'} k'(F(x)) - s_{yy}(x, F(x)) > 0. \tag{51}$$

They also imply F is continuously differentiable on $X \setminus F^{-1}(Z)$ at precisely those points x where $k'(F(x)) > s_{yy}(x, F(x))$.

Proof. The right hand side of (49) belongs to $C(X)$, so the left hand side is also continuous and bounded. Thus we deduce $\ell \|Df\|_{L^\infty(X')} \leq \|D_x s_y\|_{L^\infty(X' \times F(X'))} \leq \|s\|_{C^2(X \times Y)} < +\infty$ as desired.

Since $k'(F(x)) - s_{yy}(x, F(x))$ depends continuously on $x \in X \setminus F^{-1}(Z)$, (49) implies the same is true of $DF(x)$ provided $k'(F(x)) > s_{yy}(x, F(x))$. Conversely, non-degeneracy of $D_x s_y$ prevents DF from being locally bounded where $k'(F(x)) - s_{yy}(x, F(x))$ vanishes. ■

We now proceeed to the main result of this section, which uses condition (51) to bootstrap higher regularity for F and u from that already established for v

Corollary 23 (Regularity of optimal maps and potentials) *Suppose $X' \subset X \subset \mathbf{R}^m$, $Y' \subset Y \subset \mathbf{R}$, (s, f, g) and k satisfy the hypotheses of Theorem 16 for some $r \geq 1$. If the speed limit condition $\ell > 0$ from* (51) *holds, then the restriction of the optimal map $F : X \to Y$ to X' lies in $C^{r,1}(X')$ and $\|F\|_{C^{r,1}(X')}$ is controlled by ℓ, $\|s_y\|_{C^{r,1}(X' \times Y')}$ and $\|k\|_{C^{r,1}(Y')}$. Furthermore, if $s \in C^{r+1,1}(X' \times Y')$ then there exist minimizers (u, v) of* (5) *whose restrictions lie in $C^{r+1,1}(X') \times C^{r+1,1}(Y')$ with norms controlled by the same constants plus $\|s\|_{C^{r+1,1}(X' \times Y')}$.*

Proof. Since $k = dv/dy$, the regularity asserted for v follows directly from Theorem 16. That asserted for F then follows from the implcit function theorem after recalling that $k(F(x)) = s_y(x, F(x))$; this in turn implies that asserted for u through (50). ■

Appendices

Appendix A Self contained proof that nesting for all marginals reduces dimension

We now offer a self contained proof of Proposition 5. Our proof requires an additional definition and preliminary lemma. We say that *the level sets of $x \mapsto \frac{\partial s}{\partial y}(x, y)$ are independent of y* if for any $y_0, y_1 \in Y$, $S \subseteq X$ is a level set of $x \mapsto \frac{\partial s}{\partial y}(x, y_0)$ if and only if it is a level set of $x \mapsto \frac{\partial s}{\partial y}(x, y_1)$. In other words, for each $k_0 \in \mathbf{R}$ there exists $k_1 \in \mathbf{R}$ such that $X(y_0, k_0) = X(y_1, k_1)$, and conversely for each k_1 there exists k_0 satisfying the same conclusion.

Lemma 24 *Under the assumptions of Proposition 5, the surplus s takes pseudo-index form if and only if the level sets of $x \mapsto \frac{\partial s}{\partial y}(x, y)$ are independent of y. If s has pseudo-index form then the mixed partials $\frac{\partial^2 \sigma}{\partial I \partial y} = \frac{\partial^2 \sigma}{\partial y \partial I}$ exist and are continuous throughout $I(X) \times Y$.*

Proof. If $s(x, y) = \sigma(I(x), y) + \alpha(x)$ is of pseudo-index form, we have

$$\frac{\partial s}{\partial y}(x, y) = \frac{\partial \sigma}{\partial y}(I(x), y). \tag{52}$$

Therefore, for every y, each level set of $I(x)$ is *contained* in a level set of $x \mapsto \frac{\partial s}{\partial y}(x, y)$. If, in addition, we show that $I \mapsto \frac{\partial \sigma}{\partial y}(I, y)$ is injective, then the opposite inclusion will hold, and so the level sets of $x \mapsto \frac{\partial s}{\partial y}(x, y)$ will be exactly the level sets of $I(x)$, which are clearly independent of y.

If $\sigma \in C^2$, then

$$D_x \frac{\partial s}{\partial y}(x, y) = \frac{\partial^2 \sigma}{\partial I \partial y}(I(x), y) DI(x), \tag{53}$$

the non-degeneracy condition implies that $\frac{\partial^2 \sigma}{\partial I \partial y}$ is nowhere vanishing on $I(X) \times (a, b)$. As the continuous image of a connected set, $I(X)$ is connected, and hence an interval; we must therefore have either $\frac{\partial^2 \sigma}{\partial I \partial y} < 0$ everywhere or $\frac{\partial^2 \sigma}{\partial I \partial y} > 0$ everywhere (i.e., the Spence-Mirrlees sub- or super-modularity condition holds). Therefore,

$$I \mapsto \frac{\partial \sigma}{\partial y}(I, y)$$

is injective, and so we conclude that the level sets of $x \mapsto \frac{\partial s}{\partial y}(x, y)$ are independent of y .

If $\sigma \in C^1 \setminus C^2$, we shall use the non-degeneracy of $s \in C^2$ and identity (19) to argue that the mixed partials exist and are continuous, in which case the analysis of the preceding paragraph applies. For fixed x, non-degeneracy implies

$$D_x s(x, y) = \frac{\partial \sigma}{\partial I}(I(x), y) DI(x) \tag{54}$$

cannot vanish in any subinterval of (a, b), whence $DI(x) \neq 0$. Using the inverse function theorem (e.g. along an integral curve of the vector field DI), we deduce differentiability of $\sigma_y(I, y)$ with respect to I, and continuous dependence of the resultant derivative on both (I, y) from (52) noting $s \in C^2$, thereby justifying (53) and the resulting conclusions. On the other hand, we can also differentiate (54) with respect to y to obtain the other mixed partial, and compare the result to (53) to deduce the equality of mixed partials.

Conversely, suppose that the level sets of $x \mapsto \frac{\partial s}{\partial y}(x,y)$ are independent of y and set $I(x) = \frac{\partial s}{\partial y}(x, \frac{a+b}{2})$. We claim that, for all y,

$$x \mapsto s(x,y) - s(x, \frac{a+b}{2})$$

is constant along the level sets of $I(x)$. This will complete the proof, as then we can define unambiguously the function $\sigma(z,y) := s(x,y) - s(x, \frac{a+b}{2})$, where $x \in I^{-1}(z)$, which implies

$$s(x,y) = \sigma(I(x), y) + s(x, \frac{a+b}{2})$$

and hence $\sigma \in C^1$, noting continuity of $DI(x) = D_x s_y(x, \frac{a+b}{2}) \neq 0$.

To see the claim, let $x, \bar{x}$ lie in a level set of I, ie, $I(x) = I(\bar{x})$; we need to show $h(y) := s(x,y) - s(x, \frac{a+b}{2}) - s(\bar{x}, y) + s(\bar{x}, \frac{a+b}{2}) = 0$ for all $y \in (a,b)$. We clearly have $h(\frac{a+b}{2}) = 0$. On the other hand, differentiating yields

$$h'(y) = \frac{\partial s}{\partial y}(x,y) - \frac{\partial s}{\partial y}(\bar{x}, y).$$

Now $I(x) = I(\bar{x})$ implies $h'(\frac{a+b}{2}) = 0$. The assumed independence of the level sets therefore asserts $h'(y) = 0$ for all $y \in (a,b)$, yielding the desired result. ■

We are now ready to prove the proposition.

Proof of Proposition 5. First assume s has a pseudo-index structure; by nondegeneracy, we can assume, without loss of generality, that $\frac{\partial^2 \sigma}{\partial I \partial y} > 0$; the derivatives in question exist by Lemma 24. In this case, for any probability measures μ and ν, and $y_1 > y_0 \in Y$, the same lemma implies that sets $X_<(y_0, k^\pm(y_0))$ and $X_<(y_1, k^\pm(y_1))$ correspond exactly to sublevel sets of $I(x)$:

$$X_<(y_i, k^\pm(y_i)) = \{x : I(x) < c^\pm(y_i)\}$$

where $[c^-(y), c^+(y)]$ is the maximal interval such that $I_\# \mu[(-\infty, z)] = \nu[(-\infty, y)]$ for all $z \in [c^-(y), c^+(y)]$. As we have

$$\mu\big[X_<(y_0, k^+(y_0))\big] = \nu[(0, y_0)] \leq \nu[(0, y_1)] = \mu\big[X_<(y_1, k^-(y_1))\big] \quad (55)$$

this implies that $c^+(y_0) \leq c^-(y_1)$ and so

$$X_<(y_0, k^+(y_0)) \subset X_<(y_1, k^-(y_1)). \quad (56)$$

If, in addition, $\nu[(y_0, y_1)] > 0$, the inequality in (55) is strict, and so we must have strict containment in (56). Therefore, the model is nested.

On the other hand, assume that s does not have a pseudo-index form; we will construct measures μ and ν for which nestedness fails.

We claim that there exist y_0, y_1, k_0, k_1 such that the pairwise disjoint, open sets

$$\begin{aligned}
A_{<<} &= \{x \mid \frac{\partial s(x, y_0)}{\partial y} < k_0, \frac{\partial s(x, y_1)}{\partial y} < k_1\} \\
A_{<>} &= \{x \mid \frac{\partial s(x, y_0)}{\partial y} < k_0, \frac{\partial s(x, y_1)}{\partial y} > k_1\} \\
A_{><} &= \{x \mid \frac{\partial s(x, y_0)}{\partial y} > k_0, \frac{\partial s(x, y_1)}{\partial y} < k_1\} \\
A_{>>} &= \{x \mid \frac{\partial s(x, y_0)}{\partial y} > k_0, \frac{\partial s(x, y_1)}{\partial y} > k_1\}
\end{aligned}$$

are all nonempty. Their union is X, minus a set of Hausdorff dimension $n-1$, thanks to non-degeneracy and Proposition 2.

We first show that the claim implies the desired result, and then prove the claim. We can assume, after rescaling, that $Y = (0, 1)$. Assuming without loss of generality that $y_1 > y_0$, for some small $\epsilon > 0$, we can take μ to be a probability measure assigning the following values:

$$\begin{aligned}
\mu(A_{<<}) &= y_0 - \epsilon, \\
\mu(A_{<>}) &= \epsilon, \\
\mu(A_{><}) &= y_1 - y_0 + \epsilon, \\
\mu(A_{>>}) &= 1 - y_1 - \epsilon.
\end{aligned}$$

Taking ν to be uniform measure, the above choice of μ leads to $k(y_0) = k_0$ and $k(y_1) = k_1$. In particular, this implies that the (nonempty) set $A_{><} \subseteq X_{\leq}(y_1, k(y_1))$ but that $A_{><}$ is *disjoint* from $X_{\leq}(y_0, k(y_0))$, violating nestedness.

To see the claim, we invoke the lemma to conclude that there exists y_0, y_1 and k_0, such that the level set

$$X(y_0, k_0) = \{x \mid \frac{\partial s(x, y_0)}{\partial y} = k_0\}$$

is not a level set of $x \mapsto \frac{\partial s(x,y_1)}{\partial y}$. Choose any $\bar{x} \in X(y_0, k_0)$ and set $k_1 = \frac{\partial s(\bar{x},y_1)}{\partial y}$ and $X(y_1, k_1) := \{x \mid \frac{\partial s(x,y_1)}{\partial y} = k_1\}$. As we cannot have

$$X(y_0, k_0) = X(y_1, k_1),$$

we assume, without loss of generality, that there exists $x \in X(y_1, k_1)$ such that $\frac{\partial s(x,y_0)}{\partial y} < k_0$.

Now, if the C^1 hypersurfaces $X(y_1, k_1)$ and $X(y_0, k_0)$ interesect tranversally at $\bar{x}$, the result follows easily: in this case $X(y_1, k_1)$ must intersect both $X_<(y_0, k_0)$ and $X_>(y_0, k_0)$ and at the intersection points, one can move slightly off the set $X(y_1, k_1)$, into either of $X_<(y_1, k_1)$ or $X_>(y_1, k_1)$, and remain in $X_<(y_0, k_0)$ or $X_>(y_0, k_0)$.

If not, the intersection in tangential at $\bar{x}$. The nondegeneracy condition ensures that we may choose k_1' close to k_1 such that the set $X(y_1, k_1')$ intersects $X_>(y_0, k_0)$ near $\bar{x}$ (simply by moving a small distance from $\bar{x}$ in the direction $D_x \frac{\partial s}{\partial y}(\bar{x}, y_0)$ to obtain a new point $\bar{x}'$ and setting $k_1' = \frac{\partial s}{\partial y}(\bar{x}', y_1)$). This implies that $A_{><}$ and $A_{>>}$ are both nonempty, after replacing k_1 with k_1'. Nondegeneracy and the implicit function theorem also implies that $X(y_1, k_1')$ intersects $X_<(y_0, k_0)$ near x, which implies that $A_{<<}$ and $A_{<>}$ are both nonempty, completing the proof of the claim. ■

References

[1] Najma Ahmad, Hwa Kil Kim, and Robert J McCann. Optimal transportation, topology and uniqueness. *Bull. Math. Sci.*, 1:13–32, 2011.

[2] Luigi Ambrosio and Gianni Dal Maso. A general chain rule for distributional derivatives. *Proc. Amer. Math. Soc.*, 108(3):691–702, 1990.

[3] Gary S Becker. A theory of marriage. Part I. *J. Political Econom.*, 81:813–846, 1973.

[4] Yann Brenier. Decomposition polaire et rearrangement monotone des champs de vecteurs. *C.R. Acad. Sci. Pair. Ser. I Math.*, 305:805–808, 1987.

[5] Luis A Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5:99–104, 1992.

[6] Luis A Caffarelli. Boundary regularity of maps with convex potentials — II. *Ann. of Math. (2)*, 144:453–496, 1996.

[7] Pierre-Andre Chiappori, Robert J McCann, and Lars Nesheim. Hedonic price equilibria, stable matching and optimal transport; equivalence, topology and uniqueness. *Econom. Theory.*, 42(2):317–354, 2010.

[8] Pierre-Andre Chiappori, Robert J McCann, and Brendan Pass. Multidimensional matching. *Preprint.*

[9] Frank H Clarke. *Optimization and nonsmooth analysis*. Wiley, New York, 1983.

[10] Juan Antonio Cuesta-Albertos and Carlos Matrán. Notes on the Wasserstein metric in Hilbert spaces. *Ann. Probab.*, 17:1264–1276, 1989.

[11] Michael J Cullen. *A Mathematical Theory of Large Scale Atmosphere/Ocean Flows*. Imperial College Press, London, 2006.

[12] Michael JP Cullen and R James Purser. An extended Lagrangian model of semi-geostrophic frontogenesis. *J. Atmos. Sci.*, 41:1477–1497, 1984.

[13] Guido De Philippis and Alessio Figalli. The Monge-Ampère equation and its link to optimal transportation. *Bull. Amer. Math. Soc. (N.S.)*, 51(4):527–580, 2014.

[14] Lawrence C Evans and Ronald F Gariepy. *Measure Theory and Fine Properties of Functions*. Stud. Adv. Math. CRC Press, Boca Raton, 1992.

[15] Herbert Federer. *Geometric Measure Theory*. Springer-Verlag, New York, 1969.

[16] Alessio Figalli, Young-Heon Kim, and Robert J McCann. Hölder continuity and injectivity of optimal maps. *Arch. Rational Mech. Anal.*, 209:747–795, 2013.

[17] Uriel Frisch, Sabino Matarrese, Roya Mohayaee, and Andrei Sobolevskii. A reconstruction of the initial conditions of the universe by optimal mass transportation. *Nature*, 417:260–262, 2002.

[18] Alfred Galichon. *Optimal Transport Methods in Economics*. Princeton University Press, to appear.

[19] Wilfrid Gangbo. Habilitation thesis, Universite de Metz, available at http://people.math.gatech.edu/ gangbo/publications/habilitation.pdf, 1995.

[20] Wilfrid Gangbo and Robert J McCann. The geometry of optimal transportation. *Acta Math.*, 177:113–161, 1996.

[21] Wilfrid Gangbo and Robert J McCann. Shape recognition via Wasserstein distance. *Quart. Appl. Math.*, 58(4):705–737, 2000.

[22] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(P_2(M), W_2)$. *Methods Appl. Anal.*, 18:127–158, 2011.

[23] Victor Guillemin and Alan Pollack. *Differential Topology*. Prentice-Hall, Englewood Cliffs, NJ, 1974.

[24] Steven Haker and Allen Tannenbaum. Optimal image interpolation and optical flow. In *Multidisciplinary research in control (Santa Barbara, CA, 2002)*, volume 289 of *Lecture Notes in Control and*

Inform. Sci., pages 133–143. Springer, Berlin, 2003.

[25] Steve Hofmann, Marius Mitrea, and Michael Taylor. Singular integrals and elliptic boundary problems on regular Semmes-Kenig-Toro domains. *Int. Math. Res. Not. IMRN*, 14:2567–2865, 2010.

[26] Leonid Kantorovich. On the translocation of masses. *C.R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.

[27] Jun Kitagawa and Micah Warren. Regularity for the optimal transportation problem with euclidean distance squared cost on the embedded sphere. *SIAM J. Math. Anal.*, 44:2871–2887, 2012.

[28] Martin Knott and Cyril S Smith. On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43:39–49, 1984.

[29] Donald E Knuth. *Stable marriage and its relation to other combinatorial problems*, volume 10 of *CRM Proceedings & Lecture Notes*. American Mathematical Society, Providence, RI, 1997. An introduction to the mathematical analysis of algorithms, Translated from the French by Martin Goldstein and revised by the author.

[30] Vladimir L Levin. Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem. *Set-Valued Analysis*, 7(1):7–32, 1999.

[31] Jiakun Liu, Neil S Trudinger, and Xu-Jia Wang. Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations*, 35:165–184, 2010.

[32] Gregoire Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202:241–283, 2009.

[33] George G Lorentz. An inequality for rearrangements. *Amer. Math. Monthly*, 60:176–179, 1953.

[34] Xi-Nan. Ma, Neil S Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Rational Mech. Anal.*, 177:151–183, 2005.

[35] Robert J McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80:309–323, 1995.

[36] Robert J McCann. A convexity principle for interacting gases. *Adv. Math.*, 128:153–179, 1997.

[37] Robert J McCann. A glimpse into the differential topology and geometry of optimal transport. *Discrete Contin. Dyn. Syst.*, 34:1605–1621, 2014.

[38] Robert J McCann and Ludovic Rifford. The intrinsic dynamics of optimal transport. To appear in *J. Éc. polytech. Math.* 2016.

[39] James A Mirrlees. An exploration in the theory of optimum income taxation. *Rev. Econom. Stud.*, 38:175–208, 1971.

[40] Gaspard Monge. Mémoire sur la théorie des déblais et de remblais. *Histoire de l'Académie Royale des Sciences de Paris, avec les*

Mémoires de Mathématique et de Physique pour la même année, pages 666–704, 1781.

[41] Brendan Pass. Regularity of optimal transportation between spaces with different dimensions. *Math. Res. Lett.*, 19(2):291–307, 2012.

[42] Svetlozar T Rachev and Lüdger Rüschendorf. *Mass Transportation Problems*. Probab. Appl. Springer-Verlag, New York, 1998.

[43] Jean-Charles Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *J. Math. Econom.*, 16:191–200, 1987.

[44] Jean-Charles Rochet and Philippe Choné. Ironing, sweeping and multidimensional screening. *Econometrica*, 66:783–826, 1998.

[45] Lüdger Rüschendorf and Svetlozar T Rachev. A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.*, 32:48–54, 1990.

[46] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians. Preprint*, 2015.

[47] Lloyd S Shapley and Martin Shubik. The assignment game I: The core. *International Journal of Game Theory*, 1(1):111–130, 1971.

[48] Michael Spence. Job market signaling. *Quarterly J. Econom.*, 87:355–374, 1973.

[49] Neil S Trudinger and Xu-Jia Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8:1–32, 2009.

[50] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, 2003.

[51] Cédric Villani. *Optimal transport: old and new*, volume 338 of *Grundlehren der mathematischen Wissenschaften.* Springer, New York, 2009.